

PCT/JP03/14626

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

18.11.03

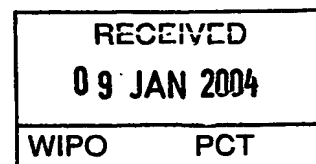
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2002年11月21日

出 願 番 号  
Application Number: 特願2002-338652  
[ST. 10/C]: [JP2002-338652]

出 願 人  
Applicant(s): 松下電器産業株式会社

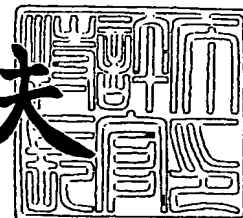


PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2003年12月18日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



出証番号 出証特2003-3105035

【書類名】 特許願

【整理番号】 2033840199

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 芳澤 伸一

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100109210

【弁理士】

【氏名又は名称】 新居 広守

【電話番号】 06-4806-7530

【手数料の表示】

【予納台帳番号】 049515

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0213583

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 標準モデル作成装置および標準モデル作成方法

【特許請求の範囲】

【請求項 1】 事象の出力確率によって定義される認識または認証用のモデルである標準モデルを作成する装置であって、

特定の対象を認識または認証するために予め作成されたモデルである参照モデルを 2 種類以上記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 2 種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と

を備えることを特徴とする標準モデル作成装置。

【請求項 2】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識または認証用のモデルである標準モデルを作成する装置であって、

特定の対象を認識または認証するために予め作成されたモデルである参照モデルを 2 種類以上記憶する参照モデル記憶手段と、

前記参照モデル記憶手段に記憶された 2 種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段と

を備えることを特徴とする標準モデル作成装置。

【請求項 3】 前記事象は、隠れマルコフモデルの状態であることを特徴とする請求項 1 または 2 記載の標準モデル作成装置。

【請求項 4】 前記標準モデル作成装置は、さらに、  
外部から参照モデルを取得して前記参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備手段を備える

ことを特徴とする請求項 1 ～ 3 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 5】 前記標準モデル作成装置は、さらに、  
認識または認証の対象に関する情報である利用情報に基づいて、前記参照モデル

ル記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する参照モデル選択手段とを備え、

前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項1～4のいずれか1項に記載の標準モデル作成装置。

【請求項6】 前記標準モデル作成装置は、さらに、

認識または認証の対象に関する情報である利用情報を作成する利用情報作成手段を備え、

前記参照モデル選択手段は、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する

ことを特徴とする請求項5に記載の標準モデル作成装置。

【請求項7】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

認識または認証の対象に関する情報である利用情報を前記端末装置から受信する利用情報受信手段を備え、

前記参照モデル選択手段は、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する

ことを特徴とする請求項5に記載の標準モデル作成装置。

【請求項8】 前記標準モデル作成手段は、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項1～7のいずれか1項に記載の標準モデル作成装置。

【請求項9】 前記標準モデル作成装置は、さらに、

作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作

成手段を備え、

前記標準モデル作成手段は、前記仕様情報作成手段が作成した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 8 記載の標準モデル作成装置。

【請求項 10】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

作成する標準モデルの仕様に関する情報である仕様情報を前記端末装置から受信する仕様情報受信手段を備え、

前記標準モデル作成手段は、前記仕様情報受信手段が受信した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 8 記載の標準モデル作成装置。

【請求項 11】 前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、

前記標準モデル作成手段は、前記仕様情報に基づいて、前記標準モデルの混合数を決定する

ことを特徴とする請求項 8 ～ 10 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 12】 前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、

前記参照モデル記憶手段は、混合数が異なる 2 以上の参照モデルを記憶し、

前記標準モデル作成手段は、混合数が異なる 2 以上の参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 1 ～ 11 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 13】 前記標準モデル作成装置は、さらに、

前記標準モデル作成手段が作成した標準モデルを記憶する標準モデル記憶手段を備える

ことを特徴とする請求項 1 ～ 12 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 14】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

前記標準モデル作成手段が作成した標準モデルを前記端末装置に送信する標準モデル送信手段を備える

ことを特徴とする請求項 1 ～ 13 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 15】 前記参照モデル準備手段は、さらに、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 4 記載の標準モデル作成装置。

【請求項 16】 前記標準モデル作成装置には、通信路を介して端末装置が接続され、

前記標準モデル作成装置は、さらに、

前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え、

前記参照モデル準備手段は、前記参照モデル受信手段が受信した参照モデルを用いて前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 15 記載の標準モデル作成装置。

【請求項 17】 前記標準モデル作成手段は、

作成する標準モデルの構造を決定する標準モデル構造決定部と、

構造が決定された前記標準モデルを特定する統計量の初期値を決定する初期標準モデル作成部と、

前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を推定して計算する統計量推定部とを有す

る

ことを特徴とする請求項 1～16 のいずれか 1 項に記載の標準モデル作成装置

。

【請求項 18】 前記標準モデル構造決定部は、前記標準モデルの構造をガウス分布または混合ガウス分布と決定し、

前記統計量推定部は、混合数が  $M$  個である混合ガウス分布により表現された前記標準モデル

【数 1】

$$\sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

(ここで、

【数 2】

$$f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

はガウス分布を表し、

【数 3】

$x$

は入力データを表す) における統計量である混合重み係数

【数 4】

$$\omega_{(m)} \quad (m = 1, 2, \dots, M)$$

、平均値

【数 5】

$$\mu_{(m)} \quad (m = 1, 2, \dots, M)$$

および分散値

【数 6】

$$\sigma_{(m)}^2 \quad (m = 1, 2, \dots, M)$$

の少なくとも1つを、前記 $N$  ( $\geq 2$ ) 種類以上の参照モデル

【数7】

$$\sum_{i=1}^{L_i} v_{(l)} g_i(x, \mu_{(l)}, \sigma_{(l)}^2) \quad (i=1, 2, \dots, N)$$

(ここで、

【数8】

$$g_i(x, \mu_{(l)}, \sigma_{(l)}^2) \quad (i=1, 2, \dots, N)$$

はガウス分布を表し、

【数9】

$$L_i \quad (i=1, 2, \dots, N)$$

は各参照モデルの混合数を表し、

【数10】

$$v_{(l)} \quad (l=1, 2, \dots, L_i)$$

は混合重み係数を表し、

【数11】

$$\mu_{(l)} \quad (l=1, 2, \dots, L_i)$$

は平均値を表し、

【数12】

$$\sigma_{(l)}^2 \quad (l=1, 2, \dots, L_i)$$

は分散値を表す) に対する前記標準モデルの尤度

【数13】

$$\log L = \sum_{i=1}^N \int_{-\infty}^{\infty} \log \left[ \sum_{m=1}^M \omega_{(m)} f(x, \mu_{(m)}, \sigma_{(m)}^2) \right] \left\{ \sum_{l=1}^{L_i} v_{(l)} g_i(x, \mu_{(l)}, \sigma_{(l)}^2) \right\} dx$$

を最大化または極大化するように、計算する



ことを特徴とする請求項 17 記載の標準モデル作成装置。

【請求項 19】 前記統計量推定部は、

【数 14】

$$\gamma(x, k) = \frac{\omega_{(k)} f(x, \mu_{(k)}, \sigma_{(k)}^2)}{\sum_{m=1}^M \omega_{(m)} f(x, \mu_{(m)}, \sigma_{(m)}^2)} \approx \frac{\omega_{(k)} f(x, \mu_{(k)}, \sigma_{(k)}^2)}{h(x, \mu, \sigma^2)}$$

$(k = 1, 2, \dots, M)$

(ここで、

【数 15】

$$h(x, \mu, \sigma^2)$$

は、

【数 16】

$$\mu$$

を平均値とし、

【数 17】

$$\sigma^2$$

を分散値とする単一のガウス分布) による近似式を利用して前記統計量を計算する

ことを特徴とする請求項 18 記載の標準モデル作成装置。

【請求項 20】 前記統計量推定部は、前記単一のガウス分布における平均値および分散値を、それぞれ、

【数 18】

$$\mu = \frac{\sum_{m=1}^M \omega_{(m)} \mu_{(m)}}{\sum_{m=1}^M \omega_{(m)}}$$

および

【数 19】

$$\sigma^2 = \frac{\sum_{m=1}^M \omega_{(m)} (\sigma_{(m)}^2 + \mu_{(m)}^2)}{\sum_{m=1}^M \omega_{(m)}} - \mu^2$$

に従って算出する第 1 近似を行う

ことを特徴とする請求項 19 記載の標準モデル作成装置。

【請求項 21】 前記統計量推定部は、前記近似式における出力分布

【数 20】

$$\omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2) \quad (m = 1, 2, \dots, M)$$

の中から、前記近似式における出力分布

【数 21】

$$\omega_{(k)} f(x; \mu_{(k)}, \sigma_{(k)}^2)$$

に距離的に近い P (≧ 1) 個の出力分布

【数 22】

$$\omega_{(p)} f(x; \mu_{(p)}, \sigma_{(p)}^2) \quad (p = 1, 2, \dots, P)$$

を選択し、選択した P 個の出力分布を用いて、前記単一のガウス分布における  
 平均値および分散値を、それぞれ、

【数 2 3】

$$\mu = \frac{\sum_{p=1}^P \omega_{(p)} \mu_{(p)}}{\sum_{p=1}^P \omega_{(p)}}$$

【数 2 4】

$$\sigma^2 = \frac{\sum_{p=1}^P \omega_{(p)} (\sigma_{(p)}^2 + \mu_{(p)}^2)}{\sum_{p=1}^P \omega_{(p)}} - \mu^2$$

に従って算出する第 2 近似を行う

ことを特徴とする請求項 19 記載の標準モデル作成装置。

【請求項 22】 前記統計量推定部は、前記第 2 近似を繰り返し、その繰り返し回数に対応して前記 P を変化させる

ことを特徴とする請求項 21 記載の標準モデル作成装置。

【請求項 23】 前記統計量推定部は、

【数 2 5】

$$\gamma(x, k) = \frac{\omega_{(k)} f(x, \mu_{(k)}, \sigma_{(k)}^2)}{\sum_{m=1}^M \omega_{(m)} f(x, \mu_{(m)}, \sigma_{(m)}^2)} \approx 1.0$$

による近似式を利用して前記統計量を計算する第 3 近似を行う

ことを特徴とする請求項 18 記載の標準モデル作成装置。

【請求項 24】 前記統計量推定部は、請求項 20 記載の第 1 近似、請求項 21 記載の第 2 近似および請求項 23 記載の第 3 近似の少なくとも 2 つを組み合わせて前記標準モデルの統計量を推定して計算する

ことを特徴とする請求項 20～23 のいずれか 1 項に記載の標準モデル作成装置。

【請求項 25】 事象の出力確率によって定義される認識または認証用のモ

デルである標準モデルを作成する方法であって、

特定の対象を認識または認証するために予め作成されたモデルである 2 種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップ

を含むことを特徴とする標準モデル作成方法。

【請求項 26】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識または認証用のモデルである標準モデルを作成する方法であって、

特定の対象を認識または認証するために予め作成されたモデルである 2 種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成ステップ

を含むことを特徴とする標準モデル作成方法。

【請求項 27】 前記事象は、隠れマルコフモデルの状態であることを特徴とする請求項 25 または 26 記載の標準モデル作成方法。

【請求項 28】 前記標準モデル作成方法は、さらに、外部から参照モデルを取得して記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備ステップと、

認識または認証の対象に関する情報である利用情報に基づいて、前記記憶手段に記憶されている参照モデルの中から 2 種類以上の参照モデルを選択する参照モデル選択ステップとを含み、

前記標準モデル作成ステップでは、前記参照モデル選択ステップで選択された参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 25～27 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 29】 前記標準モデル作成方法は、さらに、

認識または認証の対象に関する情報である利用情報を作成する利用情報作成ステップを含み、

前記参照モデル選択ステップでは、前記利用情報作成ステップで作成された前記利用情報に基づいて、前記記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する

ことを特徴とする請求項28記載の標準モデル作成方法。

【請求項30】 前記標準モデル作成方法は、さらに、

認識または認証の対象に関する情報である利用情報を外部装置から受信する利用情報受信ステップを含み、

前記参照モデル選択ステップでは、前記利用情報受信ステップで受信された前記利用情報に基づいて、前記記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する

ことを特徴とする請求項28記載の標準モデル作成方法。

【請求項31】 前記標準モデル作成ステップでは、作成する標準モデルの仕様に関する情報である仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項25～30のいずれか1項に記載の標準モデル作成方法。

【請求項32】 前記標準モデル作成方法は、さらに、作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作成ステップを含み、

前記標準モデル作成ステップでは、前記仕様情報作成ステップで作成された仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項31記載の標準モデル作成方法。

【請求項33】 前記標準モデル作成方法は、さらに、

作成する標準モデルの仕様に関する情報である仕様情報を外部装置から受信する仕様情報受信ステップを含み、

前記標準モデル作成ステップでは、前記仕様情報受信ステップが受信した仕様

情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 31 記載の標準モデル作成方法。

【請求項 34】 前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、

前記標準モデル作成ステップでは、前記仕様情報に基づいて、前記標準モデルの混合数を決定する

ことを特徴とする請求項 31～33 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 35】 前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、

前記記憶手段には、混合数が異なる 2 以上の参照モデルが記憶され、

前記標準モデル作成ステップでは、混合数が異なる 2 以上の参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算する

ことを特徴とする請求項 25～34 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 36】 前記標準モデル作成方法は、さらに、

前記標準モデル作成ステップで作成された標準モデルを外部装置に送信する標準モデル送信ステップを含む

ことを特徴とする請求項 25～35 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 37】 前記参照モデル準備ステップでは、さらに、前記記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 28 記載の標準モデル作成方法。

【請求項 38】 前記標準モデル作成方法は、さらに、

前記外部装置から送信される参照モデルを受信する参照モデル受信ステップを含み、

前記参照モデル準備ステップでは、前記参照モデル受信ステップが受信した参

照モデルを用いて前記記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行う

ことを特徴とする請求項 37 記載の標準モデル作成方法。

【請求項 39】 前記標準モデル作成ステップは、  
作成する標準モデルの構造を決定する標準モデル構造決定サブステップと、  
構造が決定された前記標準モデルを特定する統計量の初期値を決定する初期標準モデル作成サブステップと、

前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を推定して計算する統計量推定サブステップとを有する

ことを特徴とする請求項 25～38 のいずれか 1 項に記載の標準モデル作成方法。

【請求項 40】 前記標準モデル構造決定サブステップでは、前記標準モデルの構造をガウス分布または混合ガウス分布と決定し、

前記統計量推定サブステップでは、混合数が M 個である混合ガウス分布により表現された前記標準モデル

【数 26】

$$\sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

(ここで、

【数 27】

$$f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

はガウス分布を表し、

【数 28】

$x$

は入力データを表す) における統計量である混合重み係数

【数 2 9】

$$\omega_{(m)} \quad (m = 1, 2, \dots, M)$$

、平均値

【数 3 0】

$$\mu_{(m)} \quad (m = 1, 2, \dots, M)$$

および分散値

【数 3 1】

$$\sigma_{(m)}^2 \quad (m = 1, 2, \dots, M)$$

の少なくとも 1 つを、前記  $N$  ( $\geq 2$ ) 以上の参照モデル

【数 3 2】

$$\sum_{i=1}^L v_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \quad (i = 1, 2, \dots, N)$$

(ここで、

【数 3 3】

$$g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \quad (i = 1, 2, \dots, N)$$

はガウス分布を表し、

【数 3 4】

$$L_i \quad (i = 1, 2, \dots, N)$$

は各参照モデルの混合数を表し、

【数 3 5】

$$v_{(l)} \quad (l = 1, 2, \dots, L_i)$$

は混合重み係数を表し、



【数 3 6】

$$\mu_{(l)} \quad (l = 1, 2, \dots, L_i)$$

は平均値を表し、

【数 3 7】

$$\sigma_{(l)}^2 \quad (l = 1, 2, \dots, L_i)$$

は分散値を表す) に対する前記標準モデルの尤度

【数 3 8】

$$\log L = \sum_{i=1}^N \int_{-\infty}^{\infty} \log \left[ \sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2) \right] \left\{ \sum_{l=1}^{L_i} v_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx$$

を最大化または極大化するように、計算する

ことを特徴とする請求項 3 9 記載の標準モデル作成方法。

【請求項 4 1】 事象の出力確率によって定義される認識または認証用のモデルである標準モデルを作成するためのプログラムであって、

請求項 2 5、2 7～4 0 のいずれか 1 項に記載の標準モデル作成方法に含まれるステップをコンピュータに実行させる

【請求項 4 2】 事象の集合と事象または事象間の遷移の出力確率とによって定義される認識または認証用のモデルである標準モデルを作成するためのプログラムであって、

請求項 2 6～4 0 のいずれか 1 項に記載の標準モデル作成方法に含まれるステップをコンピュータに実行させる

ことを特徴とするプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解、確率モデルによるデータマイニングなどに用い

られる標準モデルの作成装置およびその方法に関する。

### 【0002】

#### 【従来の技術】

音声認識などのパターン認識の分野で、認識用の標準モデルとして確率モデルを用いる方法が近年注目されており、特に、隠れマルコフモデル（以下HMMと呼ぶ）や混合ガウス分布モデル（以下GMMと呼ぶ）が広く用いられている。また、意図理解において、意図、知識、嗜好などを表す標準モデルとして確率モデルを用いる方法が近年注目されており、特に、ペイジアンネットが広く用いられている。また、データマイニングの分野で、データを分類するために各カテゴリの代表モデルとして確率モデルを用いる方法が注目されており、GMMなどが広く用いられている。また、音声認証、指紋認証、虹彩認証などの認証の分野で、認証用の標準モデルとして確率モデルを用いる方法が注目されており、GMMなどが用いられている。HMMにより表現される標準モデルの学習アルゴリズムとしてバウム・ウェルチ（B a u m - W e l c h）の再推定の方法が広く用いられている（例えば、今井聖著、“音声認識”、p p . 1 5 0 - 1 5 2、共立出版株式会社、1995年11月25日発行参照）。また、GMMにより表現される標準モデルの学習アルゴリズムとしてEM（E x p e c t a t i o n - M a x i m i z a t i o n）アルゴリズムが広く用いられている（例えば、古井貞▲ひろ▼著、“音声情報処理”、p p . 1 0 0 - 1 0 4、森北出版株式会社、1998年6月30日発行参照）。バウム・ウェルチの再推定の方法、EMアルゴリズムのいずれの学習アルゴリズムも、入力データに対する確率（尤度）を最大化もしくは極大化するように標準モデルのパラメータ（統計量）を計算して標準モデルを作成する。

### 【0003】

上記の学習方法を音声認識の標準モデルの作成に用いた場合、多様な話者や雑音などの音響的特徴量の変動に対応するために多数の音声データで標準モデルを学習することが望ましい。また、意図理解に用いた場合、多様な話者や状況などの変動に対応するために多数のデータで標準モデルを学習することが望ましい。また、虹彩認証に用いた場合、太陽光、カメラ位置・回転などの変動に対応する

ために多数の虹彩画像データで標準モデルを学習することが望ましい。しかしながら、このような多量のデータを取り扱う場合、学習の収束までに時間がかかり、また収束値も全体の最適値ではなく局所的な最適値になる可能性が高い。

#### 【0004】

上記の課題を解決するために、ガウス分布で表現された複数のモデル（以下、標準モデルの作成のために参照用として準備されるモデルを「参照モデル」と呼ぶ。）を準備し、準備した参照モデルを統合して標準モデルを作成する方法が提案されている。参照モデルごとに学習を行うため、学習の収束が早くなり、また収束値も全体の最適値なる可能性が高い。準備された複数の参照モデルはさまざまな方法により統合され標準モデルが作成される。

#### 【0005】

第1の従来方法では、準備した参照モデルはGMMで表現されており、複数の参照モデルのGMMを重み付きで混合することで標準モデルを作成している（例えば、特許文献1参照）。

#### 【0006】

また、第2の従来方法では、第1の従来方式に加えて、入力データに対する確率（尤度）を最大化あるいは極大化して線形結合された混合重みを学習することで標準モデルを作成している（例えば、特許文献2参照）。この第2の従来方法により、第1の従来方法よりも高精度の標準モデルが作成できる。また、入力データとして特定の話者のデータを用いたときには標準モデルとして話者適応モデルを作成できる。複数の参照モデルを事前に作成しておけば、短時間に適応モデルを作成することができる。

#### 【0007】

また、第3の従来方法では、標準モデルの平均値を準備した参照モデルの平均値の線形結合で表現し、入力データに対する確率（尤度）を最大化あるいは極大化して線形結合係数を学習することで標準モデルを作成している。ここでは入力データとして特定話者の音声データを用いており標準モデルを話者適応モデルとして用いている（例えば、非特許文献1参照）。複数の参照モデルを事前に作成しておけば、短時間に適応モデルを作成することができる。

## 【0008】

また、第4の従来方法では、準備した参照モデルは単一ガウス分布で表現されており、複数の参照モデルのガウス分布を混合したのちに、クラスタリングにより同一クラスに属するガウス分布を統合することで標準モデルを作成している（例えば、特許文献3参照）。これによって、標準モデルのガウス分布の混合数を制御することができる。

## 【0009】

また、第5の従来方法では、複数の参照モデルは同数の混合数の混合ガウス分布で表現され、各ガウス分布には通し番号が付与されている。標準モデルは、同一の通し番号をもつガウス分布を統合することにより作成される。統合する複数の参照モデルは利用者に音響的に近い話者で作成されたモデルであり、作成させる標準モデルは話者適応モデルである（例えば、非特許文献2参照）。複数の参照モデルを事前に作成しておけば、短時間に適応モデルを作成することができる。

## 【0010】

## 【特許文献1】

特開平4-125599号公報（第3項、第1図）

## 【0011】

## 【特許文献2】

特開平10-268893号公報（第3-6頁、第1図）

## 【0012】

## 【特許文献3】

特開平9-81178号公報（第3-4項、第4図、第5図）

## 【0013】

## 【非特許文献1】

M. J. F. Gales, "Cluster Adaptive Training For Speech Recognition", 1998年、ICSLP98予稿集、pp. 1783-1786

## 【0014】

## 【非特許文献2】

芳澤、外6名、"十分統計量と話者距離を用いた音韻モデルの教師なし学習法"、2002年3月1日、電子情報通信学会、Vol. J85-D-II、No. 3、pp. 382-389

## 【0015】

## 【発明が解決しようとする課題】

しかしながら、第1の従来方法では、統合する参照モデル数の増加とともに標準モデルの混合数が増加して、標準モデルのための記憶容量、認識処理量が膨大となり実用的でない。また、標準モデルの混合数を制御することができない。また、標準モデルは準備した参照モデルの単純な混合和であるため、高精度の標準モデルが作成できない。

## 【0016】

第2の従来方法では、統合する参照モデル数の増加とともに標準モデルの混合数が増加して、標準モデルのための記憶容量、認識処理量が膨大となり実用的でない。また、標準モデルの混合数を制御することができない。また、標準モデルは、準備した参照モデルの単純な混合和であり学習するパラメータが混合重みに限定されているため、高精度の標準モデルが作成できない。

## 【0017】

第3の従来方法では、学習するパラメータが準備した参照モデルの平均値の線形結合係数に限定されているため高精度の標準モデルが作成できない。

第4の従来方法では、クラスタリングをヒューリスティックに行うため高精度の標準モデルを作成することが困難である。また、準備する参照モデルは単一のガウス分布であるため精度が低く、それらを統合した標準モデルの精度は低い。

## 【0018】

第5の従来方法では、標準モデルは、同一の通し番号をもつガウス分布を統合することにより作成されるが、最適な標準モデルを作成するためには、一般的には統合するガウス分布は1対1に対応するとは限らない。また、複数の参照モデルが異なる混合数をもつ場合に標準モデルを作成することができない。また、標準モデルの混合数を制御することができない。また、参照モデルにおけるガウス

分布に通し番号が付与されていない場合に標準モデルを作成することができない。

#### 【0019】

そこで、本発明は、このような問題点に鑑みてなされたものであり、隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解、確率モデルによるデータマイニングなどに用いられる高精度な標準モデルを作成する標準モデル作成装置等を提供することを目的とする。

#### 【0020】

また、本発明は、学習のための入力データや教師データを必要とすることなく、簡易に標準モデルを作成することが可能な標準モデル作成装置等を提供することをも目的とする。

さらに、本発明は、標準モデルを利用する認識や認証の対象にふさわしい標準モデルを作成したり、標準モデルを用いて認識や認証処理を実行する装置の仕様や環境に適した標準モデルを作成することが可能な汎用性および柔軟性に優れた標準モデル作成装置等を提供することをも目的とする。

#### 【0021】

##### 【課題を解決するための手段】

上記目的を達成するために、本発明に係る標準モデル作成装置は、事象の集合の出力確率によって定義される認識または認証用のモデルである標準モデルを作成する装置であって、特定の対象を認識または認証するために予め作成されたモデルである参照モデルを2種類以上記憶する参照モデル記憶手段と、前記参照モデル記憶手段に記憶された2種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように当該標準モデルの統計量を計算することによって標準モデルを作成する標準モデル作成手段とを備えることを特徴とする。

#### 【0022】

これによって、2種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように標準モデルの統計量が計算され、標準モデルが

作成されるので、音声データ等の入力データや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した高精度な標準モデルが作成される。

#### 【0023】

ここで、前記標準モデル作成装置は、さらに、外部から参照モデルを取得して前記参照モデル記憶手段に格納すること、および、参照モデルを作成して前記参照モデル記憶手段に格納することの少なくとも一方を行う参照モデル準備手段を備えてもよい。これによって、標準モデル作成装置の外部から新たな参照モデルを取り込み、取り込んだ参照モデルに基づいた標準モデルの作成が可能となるので、様々な認識・認証対象に対応した汎用性の高い標準モデル作成装置が実現される。

#### 【0024】

また、前記標準モデル作成装置は、さらに、認識または認証の対象に関する情報である利用情報を作成する利用情報作成手段と、作成された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から2種類以上の参照モデルを選択する参照モデル選択手段とを備え、前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

#### 【0025】

これによって、利用者の特徴、利用者の年齢、性別、利用環境などの利用情報に基づいて、準備された複数の参照モデルの中から認識・認証対象に適した参照モデルだけが選択され、それら参照モデルを統合した標準モデルが作成されるので、認識・認証対象により特化した精度の高い標準モデルが作成される。

#### 【0026】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、認識または認証の対象に関する情報である利用情報を前記端末装置から受信する利用情報受信手段と、受信された前記利用情報に基づいて、前記参照モデル記憶手段に記憶されている参照モデルの中から2

種類以上の参照モデルを選択する参照モデル選択手段とを備え、前記標準モデル作成手段は、前記参照モデル選択手段が選択した参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

#### 【0027】

これによって、通信路を介して送信されてきた参照モデルに基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識・認証システムの構築が実現される。

#### 【0028】

また、前記標準モデル作成装置は、さらに、作成する標準モデルの仕様に関する情報である仕様情報を作成する仕様情報作成手段を備え、前記標準モデル作成手段は、前記仕様情報作成手段が作成した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

#### 【0029】

これによって、標準モデルを使用する装置のCPUパワー、記憶容量、要求される認識精度、要求される認識処理時間などの仕様情報に基づいて標準モデルが作成されるので、特定の仕様条件を満たす標準モデルの生成が可能となり、計算エンジン等の認識・認証処理に必要なリソース環境に適した標準モデルの生成が実現される。

#### 【0030】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、作成する標準モデルの仕様に関する情報である仕様情報を前記端末装置から受信する仕様情報受信手段を備え、前記標準モデル作成手段は、前記仕様情報受信手段が受信した仕様情報に基づいて、前記参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。

#### 【0031】

これによって、通信路を介して送信されてきた仕様情報に基づいて標準モデル



が作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識・認証システムの構築が実現される。

#### 【0032】

たとえば、前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、前記標準モデル作成手段は、前記仕様情報に基づいて、前記標準モデルの混合数を決定してもよい。これによって、作成される標準モデルに含まれるガウス分布の混合数が動的に決定されることとなり、認識・認証処理が実行される環境や要求仕様等に応じて標準モデルの構造を制御することが可能となる。例として、標準モデルを使用する認識装置のCPUパワーが小さい場合、記憶容量が小さい場合、要求される認識処理時間が短い場合などは標準モデルの混合数を少なく設定して仕様に合わせることができ、一方、要求される認識精度が高い場合などは混合数を多く設定して認識精度を高くすることができる。

#### 【0033】

また、前記参照モデルおよび前記標準モデルは、ガウス分布または混合ガウス分布を用いて表現され、前記参照モデル記憶手段は、混合数が異なる2以上の参照モデルを記憶し、前記標準モデル作成手段は、混合数が異なる2以上の参照モデルに対する前記標準モデルの確率または尤度を最大化または極大化するように前記標準モデルの統計量を計算してもよい。これによって、混合数が異なる参照モデルに基づいて標準モデルが作成されるので、予め準備された多種多様な構造の参照モデルに基づく標準モデルの作成が可能となり、より認識・認証対象に適した精度の高い標準モデルの作成が実現される。

#### 【0034】

また、前記標準モデル作成装置は、さらに、前記標準モデル作成手段が作成した標準モデルを記憶する標準モデル記憶手段を備えてもよい。これによって、作成された標準モデルを一時的にバッファリングしておき、送信要求に対してすぐに出力したり、他の装置に提供するデータサーバとしての役割を果たしたりすることが可能となる。

#### 【0035】

また、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、前記標準モデル作成手段が作成した標準モデルを前記端末装置に送信する標準モデル送信手段を備えてもよい。これによって、作成された標準モデルは空間的に離れた場所に置かれた外部装置に送信されるので、本標準モデル作成装置を標準モデル作成エンジンとして独立させたり、標準モデル作成装置を通信システムにおけるサーバとして機能させたりすることが可能になる。

#### 【0036】

また、前記参照モデル準備手段は、さらに、前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行ってもよい。たとえば、前記標準モデル作成装置には、通信路を介して端末装置が接続され、前記標準モデル作成装置は、さらに、前記端末装置から送信される参照モデルを受信する参照モデル受信手段を備え、前記参照モデル準備手段は、前記参照モデル受信手段が受信した参照モデルを用いて前記参照モデル記憶手段が記憶する参照モデルの更新および追加の少なくとも一方を行ってもよい。

#### 【0037】

これによって、準備される参照モデルの追加、更新等が行われるので、様々な認識・認証対象用のモデルを参照モデルとして追加したり、より精度の高い参照モデルに置き換えたりすることが可能となり、更新した参照モデルによる標準モデルの再生成や、生成された標準モデルを参照モデルとして再び標準モデルを作成するというフィードバックによる学習等が可能となる。

#### 【0038】

なお、本発明は、このような標準モデル作成装置として実現することができるだけでなく、標準モデル作成装置が備える特徴的な構成要素をステップとする標準モデル作成方法として実現したり、それらのステップをコンピュータに実行させるプログラムとして実現したりすることができる。そして、そのプログラムをCD-ROM等の記録媒体やインターネット等の伝送媒体を介して配信することができるのは言うまでもない。

#### 【0039】

**【発明の実施の形態】**

以下、本発明の実施の形態について図面を参照しながら詳しく説明する。なお、図中同一または相当部分には同一符号を付し、その説明は繰り返さない。

**【0040】****(第1の実施の形態)**

図1は、本発明の第1の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ101に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデルを作成する場合を例にして説明する。

**【0041】**

サーバ101は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率で表現された隠れマルコフモデルによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部111と、参照モデル準備部102と、参照モデル記憶部103と、標準モデル作成部104と、書き込み部112とを備える。

**【0042】**

読み込み部111は、CD-ROMなどのストレージデバイスに書き込まれた子供用参照モデル、成人用参照モデル、高齢者用参照モデルを読み込む。参照モデル準備部102は、読み込まれた参照モデル121を参照モデル記憶部103へ送信する。参照モデル記憶部103は、3個の参照モデル121を記憶する。

**【0043】**

標準モデル作成部104は、参照モデル記憶部103が記憶した3個( $N=3$ )の参照モデル121に対する確率または尤度を最大化または極大化するように標準モデル122を作成する処理部であり、標準モデルの構造(ガウス分布の混合数など)を決定する標準モデル構造決定部104aと、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部104bと、決定された初期標準モデルを記憶する統計量記憶部104cと、統計量記憶部104cに記憶された初期標準モデルに対して、第1近似部104eによる近似計算等を用いることにより、参照モデル記憶部103に記憶

されている3個 ( $N=3$ ) の参照モデル121に対する確率または尤度を最大化または極大化するような統計量を算出する(最終的な標準モデルを生成する) 統計量推定部104dとからなる。なお、統計量とは、標準モデルを特定するパラメータであり、ここでは、混合重み係数、平均値、分散値である。

#### 【0044】

書き込み部112は、標準モデル作成部104が作成した標準モデル122をCD-ROMなどのストレージデバイスに書き込む。

次に、以上のように構成されたサーバ101の動作について説明する。

#### 【0045】

図2は、サーバ101の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する(ステップS100)。つまり、読み込み部111は、CD-ROMなどのストレージデバイスに書き込まれた子供用参照モデル、成人用参照モデル、高齢者用参照モデルを読み込み、参照モデル準備部102は、読み込まれた参照モデル121を参照モデル記憶部103へ送信し、参照モデル記憶部103は、3個の参照モデル121を記憶する。

#### 【0046】

参照モデル121は、音素ごとのHMMにより構成される。参照モデル121の一例を図3に示す。ここでは、子供用参照モデル、成人用参照モデル、高齢者用参照モデルのイメージ図が示されている(なお、本図では、高齢者用参照モデルのイメージ図は省略されている)。これら3つの参照モデルの全てが、状態数3個、各状態は混合数が3個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として12次元( $J=12$ )のケプストラム係数が用いられる。

#### 【0047】

次に、標準モデル作成部104は、参照モデル記憶部103が記憶した3個の参照モデル121に対する確率または尤度を最大化または極大化するように標準モデル122を作成する(ステップS101)。

#### 【0048】

最後に、書き込み部112は、標準モデル作成部104が作成した標準モデル

122をCD-ROMなどのストレージデバイスに書き込む（ステップS102）。CD-ROMなどのストレージデバイスに書き込まれた標準モデルは、子供、成人、高齢者を考慮した音声認識用の標準モデルとして利用される。

#### 【0049】

図4は、図2におけるステップS101（標準モデルの作成）の詳細な手順を示すフローチャートである。

まず、標準モデル構造決定部104aは、標準モデルの構造を決定する（ステップS102a）。ここでは、標準モデルの構造として、音素ごとのHMMにより構成され、3状態であり、各状態における出力分布の混合数を3個（ $M=3$ ）と決定する。

#### 【0050】

次に、初期標準モデル作成部104bは、標準モデルを計算するための統計量の初期値を決定する（ステップS102b）。ここでは、参照モデル記憶部103に記憶された3つの参照モデルを、統計処理計算を用いて1つのガウス分布に統合したものを統計量の初期値とし、その初期値を初期標準モデルとして統計量記憶部306cに記憶する。

#### 【0051】

具体的には、初期標準モデル作成部104bは、上記3つの状態I（ $I=1, 2, 3$ ）それぞれについて、以下の数39に示される出力分布を生成する。なお、式中のM（ガウス分布の混合数）は、ここでは、3である。

#### 【0052】

##### 【数39】

$$\sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

ここで、

##### 【数40】

$$f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

は、ガウス分布を表し、

【数 4 1】

 $x$ 

は、1 2 次元 ( $J = 1 2$ ) の L P C ケプストラム係数を表し、

【数 4 2】

$$\omega_{(m)} \quad (m = 1, 2, \dots, M)$$

は、各ガウス分布の混合重み係数を表し、

【数 4 3】

$$\mu_{(m)} \quad (m = 1, 2, \dots, M)$$

は、各ガウス分布の平均値を表し、

【数 4 4】

$$\sigma_{(m)}^2 \quad (m = 1, 2, \dots, M)$$

は、各ガウス分布の分散値を表す。

【0 0 5 3】

そして、統計量推定部 1 0 4 d は、参照モデル記憶部 1 0 3 に記憶された 3 つの参照モデル 1 2 1 を用いて、統計量記憶部 1 0 4 c に記憶された標準モデルの統計量を推定する (ステップ S 1 0 2 c)。

【0 0 5 4】

具体的には、3 つ ( $N = 3$ ) の参照モデル 1 2 1 の各状態  $I$  ( $I = 1, 2, 3$ ) における出力分布、即ち、以下の数 4 5 に示される出力分布に対する標準モデルの確率または尤度 (以下の数 5 1 に示される尤度  $\log L$ ) を極大化もしくは最大化するような標準モデルの統計量 (上記数 4 2 に示される混合重み係数、上記数 4 3 に示される平均値、および、上記数 4 4 に示される分散値) を推定する。

【0 0 5 5】

【数 4 5】

$$\sum_{i=1}^L v_{(i)} g_i(x; \mu_{(i)}, \sigma_{(i)}^2) \quad (i = 1, 2, \dots, N)$$

ここで、

【数 4 6】

$$g_i(x; \mu_{(i)}, \sigma_{(i)}^2) \quad (i = 1, 2, \dots, N)$$

はガウス分布を表し、

【数 4 7】

$$L_i \quad (i = 1, 2, \dots, N)$$

は各参照モデルの混合数（ここでは、3）を表し、

【数 4 8】

$$v_{(l)} \quad (l = 1, 2, \dots, L_i)$$

は各ガウス分布の混合重み係数を表し、

【数 4 9】

$$\mu_{(l)} \quad (l = 1, 2, \dots, L_i)$$

は各ガウス分布の平均値を表し、

【数 5 0】

$$\sigma_{(l)}^2 \quad (l = 1, 2, \dots, L_i)$$

は各ガウス分布の分散値を表す。

【0056】

【数 5 1】

$$\log L = \sum_{i=1}^N \int_{-\infty}^{\infty} \log \left[ \sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2) \right] \left\{ \sum_{l=1}^{L_i} v_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx$$

そして、以下の数 5 2、数 5 3 および数 5 4 に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

【0057】

【数 5 2】

$$\omega_{(k)} = \frac{\sum_{i=1}^N \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}{\sum_{i=1}^N \sum_{m=1}^M \int_{-\infty}^{\infty} \gamma(x, m) \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}$$

$(k = 1, 2, \dots, M)$

【0 0 5 8】

【数 5 3】

$$\mu_{(k,j)} = \frac{\sum_{i=1}^N \int_{-\infty}^{\infty} \gamma(x, k) x_{(j)} \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}{\sum_{i=1}^N \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}$$

$(k = 1, 2, \dots, M, j = 1, 2, \dots, J)$

【0 0 5 9】

【数 5 4】

$$\sigma_{(k,j)}^2 = \frac{\sum_{i=1}^N \int_{-\infty}^{\infty} \gamma(x, k) (x_{(j)} - \mu_{(k)})^2 \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}{\sum_{i=1}^N \int_{-\infty}^{\infty} \gamma(x, k) \left\{ \sum_{l=1}^{L_l} \nu_{(l)} g_i(x; \mu_{(l)}, \sigma_{(l)}^2) \right\} dx}$$

$(k = 1, 2, \dots, M, j = 1, 2, \dots, J)$

このとき、統計量推定部 104 d の第 1 近似部 104 e により、以下の数 5 5 に示される近似式が用いられる。

【0 0 6 0】

【数 5 5】

$$\gamma(x, k) = \frac{\omega_{(k)} f(x; \mu_{(k)}, \sigma_{(k)}^2)}{\sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2)} \approx \frac{\omega_{(k)} f(x; \mu_{(k)}, \sigma_{(k)}^2)}{h(x; \mu, \sigma^2)}$$

$(k = 1, 2, \dots, M)$



ここで、

【0061】

【数56】

$$h(x; \mu, \sigma^2)$$

は、

【0062】

【数57】

$$\mu$$

を平均値とし、

【0063】

【数58】

$$\sigma^2$$

を分散値とする単一のガウス分布を表す。

【0064】

また、第1近似部104eは、上記数56に示された単一ガウス分布の平均値（数57）および分散値（数58）を、それぞれ、以下の数59および数60に示された式に従って算出する。

【0065】

【数59】

$$\mu = \frac{\sum_{m=1}^M \omega_{(m)} \mu_{(m)}}{\sum_{m=1}^M \omega_{(m)}}$$

【0066】

【数 6 0】

$$\sigma^2 = \frac{\sum_{m=1}^M \omega_{(m)} (\sigma_{(m)}^2 + \mu_{(m)}^2)}{\sum_{m=1}^M \omega_{(m)}} - \mu^2$$

図 5 は、第 1 近似部 104 e による近似計算を説明する図である。第 1 近似部 104 e は、本図に示されるように、上記数 55 に示された近似式における単一ガウス分布（数 56）を、標準モデルを構成する全ての混合ガウス分布を用いて決定している。

【0067】

以上の第 1 近似部 104 e による近似式を考慮してまとめると、統計量推定部 104 d での計算式は次の通りになる。つまり、統計量推定部 104 d は、以下の数 61、数 62 および数 63 に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、統計量記憶部 104 c に記憶する。そして、このような統計量の推定と統計量記憶部 104 c への記憶を R（ $\geq 1$ ）回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル 122 の統計量として出力する。

【0068】

【数 6 1】

$$\omega_{(k)} = \frac{\sum_{i=1}^N \prod_{j=1}^J \sum_{l=1}^{L_i} A_{(k,l,i,j)}}{\sum_{i=1}^N \sum_{k=1}^M \prod_{j=1}^J \sum_{l=1}^{L_i} A_{(k,l,i,j)}}$$

$$A_{(k,l,i,j)} = \frac{\omega_{(k)} v_{(l,i)} \sigma_{(j)}^2}{\sqrt{\sigma_{(k,j)}^2 \sigma_{(j)}^2 + \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 - \sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2}}$$

$$\times \exp \left\{ \frac{1}{2} \left( \frac{\sigma_{(k,j)} \sigma_{(l,i,j)}}{\sigma_{(j)}^2} \mu_{(j)} - \frac{\sigma_{(k,j)} \sigma_{(j)}}{\sigma_{(l,i,j)}} \mu_{(l,i,j)} - \frac{\sigma_{(l,i,j)} \sigma_{(j)}}{\sigma_{(k,j)}} \mu_{(k,j)} \right)^2 \right. \\ \left. + \frac{\mu_{(j)}^2}{\sigma_{(j)}^2} - \frac{\mu_{(l,i,j)}^2}{\sigma_{(l,i,j)}^2} - \frac{\mu_{(k,j)}^2}{\sigma_{(k,j)}^2} \right\}$$

$(k = 1, 2, \dots, M)$

【0069】

【数 6 2】

$$\mu_{(k,j)} = \frac{\sum_{i=1}^N \sum_{l=1}^{L_i} E_{(k,l,i,j)}}{\sum_{i=1}^N \prod_{j=1}^J \sum_{l=1}^{L_i} A_{(k,l,i,j)}}$$

$$E_{(k,l,i,j)} = \frac{\sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2 \mu_{(j)} - \sigma_{(k,j)}^2 \sigma_{(j)}^2 \mu_{(l,i,j)} - \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 \mu_{(k,j)}}{\sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2 - \sigma_{(k,j)}^2 \sigma_{(j)}^2 - \sigma_{(l,i,j)}^2 \sigma_{(j)}^2} \times A_{(k,l,i,j)}$$

$$(k = 1, 2, \dots, M, j = 1, 2, \dots, J)$$

【0070】

【数 6 3】

$$\sigma_{(k,j)}^2 = \frac{\sum_{i=1}^N \sum_{l=1}^{L_l} G_{(k,l,i,j)}}{\sum_{i=1}^N \prod_{j=1}^J \sum_{l=1}^{L_l} A_{(k,l,i,j)}}$$

$$G_{(k,l,i,j)} = \left\{ \frac{\sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2 \sigma_{(j)}^2}{\sigma_{(k,j)}^2 \sigma_{(j)}^2 + \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 - \sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2} \right. \\ \left. + \left( \mu_{(k,j)} - \frac{\sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2 \mu_{(j)} - \sigma_{(k,j)}^2 \sigma_{(j)}^2 \mu_{(l,i,j)} - \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 \mu_{(k,j)}}{\sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2 - \sigma_{(k,j)}^2 \sigma_{(j)}^2 - \sigma_{(l,i,j)}^2 \sigma_{(j)}^2} \right)^2 \right\} \times A_{(k,l,i,j)}$$

なお、状態遷移確率については、HMMの対応する状態遷移確率を参照モデル 1 2 1 に対して全て加えあわせた全体が 1 になるように正規化したものを用いる。

【0 0 7 1】

以上説明したように、本発明の第 1 の実施の形態によれば、予め準備された参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、学習のための入力データや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した高精度な標準モデルが作成される。

【0 0 7 2】

なお、標準モデル 1 2 2 は、音素ごとに HMM を構成するに限らず、文脈依存の HMM で構成してもよい。

また、標準モデル 1 2 2 を構成する HMM は、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる混合数の混合ガウス分布により構成してもよい。

【0 0 7 3】

また、参照モデル 1 2 1 は、子供用参照モデル、成人用参照モデル、高齢者用参照モデルにおいて、異なる状態数により構成してもよいし、異なる混合数の混

合ガウス分布により構成してもよい。

【0074】

また、標準モデル122を用いて、サーバ101において音声認識を行ってもよい。

また、参照モデル121をCD-ROM、DVD-RAMなどのストレージデバイスから読み込む代わりに、サーバ101において音声データから参照モデル121を作成してもよい。

【0075】

また、参照モデル準備部102は、必要に応じてCD-ROM、DVD-RAMなどのストレージデバイスから読み込まれた新たな参照モデルを参照モデル記憶部103に追加・更新してもよい。つまり、参照モデル準備部102は、新たな参照モデルを参照モデル記憶部103に格納するだけでなく、同一の認識・認証対象についての参照モデルが参照モデル記憶部103に格納されている場合には、その参照モデルと置き換えることによって参照モデルを更新したり、参照モデル記憶部103に格納されている不要な参照モデルを削除してもよい。

【0076】

また、参照モデル準備部102は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部103に追加・更新してもよい。

【0077】

(第2の実施の形態)

図6は、本発明の第2の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がセットトップボックス201（以下、STBと呼ぶ）に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（話者適応モデル）を作成する場合を例にして説明する。

【0078】

STB201は、ユーザの発話を認識してTV番組の自動切替等を行うデジタル放送用受信機であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、

マイク 211 と、音声データ蓄積部 212 と、参照モデル準備部 202 と、参照モデル記憶部 203 と、利用情報作成部 204 と、参照モデル選択部 205 と、標準モデル作成部 206 と、音声認識部 213 とを備える。

#### 【0079】

マイク 211 に収集された音声データは、音声データ蓄積部 212 に蓄積される。参照モデル準備部 202 は、音声データ蓄積部 212 が蓄積した音声データを用いて話者ごとに参照モデル 221 を作成し、参照モデル記憶部 203 に記憶する。

#### 【0080】

利用情報作成部 204 は、利用情報 224 である利用者の音声をマイク 211 により収集する。ここで、利用情報とは、認識、識別、認証の対象（人・物）に関する情報であり、ここでは、音声認識の対象となる利用者の音声である。参照モデル選択部 205 は、利用情報作成部 204 が作成した利用情報 224 に基づいて、参照モデル記憶部 203 が記憶している参照モデル 221 の中から、利用情報 224 が示す利用者の音声に音響的に近い参照モデル 223 を選択する。

#### 【0081】

標準モデル作成部 206 は、参照モデル選択部 205 が選択した話者の参照モデル 223 に対する確率または尤度を最大化または極大化するように標準モデル 222 を作成する処理部であり、標準モデルの構造（ガウス分布の混合数など）を決定する標準モデル構造決定部 206a と、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部 206b と、決定された初期標準モデルを記憶する統計量記憶部 206c と、統計量記憶部 206c に記憶された初期標準モデルに対して、第 2 近似部 206e による近似計算等を用いることにより、参照モデル選択部 205 が選択した参照モデル 223 に対する確率または尤度を最大化または極大化するような統計量を算出する（最終的な標準モデルを生成する）統計量推定部 206d とからなる。

#### 【0082】

音声認識部 213 は、標準モデル作成部 206 によって作成された標準モデル 222 を用いて利用者の音声を認識する。

次に、以上のように構成されたSTB201の動作について説明する。

#### 【0083】

図7は、STB201の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（ステップS200）。つまり、マイク211によりAさんからZさんの音声データを収集して音声データ蓄積部212に蓄積する。参照モデル準備部202は、音声データ蓄積部212が蓄積した音声データを用いて話者ごとに参照モデル221をバウム・ウェルチの再推定の方法により作成する。

#### 【0084】

参照モデル記憶部203は、参照モデル準備部202が作成した参照モデル221を記憶する。参照モデル221は、音素ごとのHMMにより構成される。参照モデル221の一例を図8に示す。ここでは、AさんからZさんの全ての参照モデルが、状態数3個、各状態は混合数が5個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として25次元（ $J=25$ ）のメルケプストラム係数が用いられる。

#### 【0085】

次に、利用情報作成部204は、利用情報224である利用者の音声をマイク211により収集する（ステップS201）。参照モデル選択部205は、その利用者の音声に音響的に近い参照モデル223を、参照モデル記憶部203が記憶している参照モデル221の中から選択する（ステップS202）。具体的には、利用者の音声をAさんからZさんの参照モデルに入力して発声単語に対する尤度大きい10人（ $N=10$ ）の話者の参照モデルを選択する。

#### 【0086】

そして、標準モデル作成部206は、参照モデル選択部205が選択した10個の参照モデル223に対する確率または尤度を最大化または極大化するように標準モデル222を作成する（ステップS203）。

#### 【0087】

最後に、音声認識部213は、マイク211から介して送られてくる利用者の音声を入力とし、標準モデル作成部206で作成された標準モデル222を用い

て音声認識を行う（S204）。たとえば、利用者が発話した音声を音響解析等を行うことで25次元のメルケプストラム係数を算出し、音素ごとの標準モデル222に入力することで、高い尤度を有する音素の連なりを特定する。そして、その音素の連なりと予め受信している電子番組データ中の番組名とを比較し、一定以上の尤度が検出された場合に、その番組に切り替えるという自動番組切替の制御を行う。

#### 【0088】

次に、図7におけるステップS203（標準モデルの作成）の詳細な手順を説明する。手順の流れは、図4に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

#### 【0089】

まず、標準モデル構造決定部206aは、標準モデルの構造を決定する（図4のステップS102a）。ここでは、標準モデルの構造として、音素ごとのHMMにより構成され、3状態であり、各状態における出力分布の混合数が16個（ $M=16$ ）と決定する。

#### 【0090】

次に、初期標準モデル作成部206bは、標準モデルを計算するための統計量の初期値を決定する（図4のステップS102b）。ここでは、参照モデル選択部205が選択した10個の参照モデル223を、統計処理計算を用いて1つのガウス分布に統合したものを統計量の初期値とし、その初期値を初期標準モデルとして統計量記憶部206cに記憶する。ここでは、話者ごとに学習した5混合の参照モデルを用いて高精度な16混合の標準モデル（話者適応モデル）を作成する。

#### 【0091】

具体的には、初期標準モデル作成部206bは、上記3つの状態I（ $I=1, 2, 3$ ）それぞれについて、上記数39に示される出力分布を生成する。

ただし、本実施の形態では、上記数39に示された出力分布における

#### 【0092】



## 【数 6 4】

 $x$ 

は、25次元 ( $J=25$ ) のメルケプストラム係数を表す。

## 【0093】

そして、統計量推定部 206 d は、参照モデル選択部 205 が選択した 10 個の参照モデル 223 を用いて、統計量記憶部 206 c に記憶された標準モデルの統計量を推定する (図 4 のステップ S102 c)。

## 【0094】

つまり、10 個 ( $N=10$ ) の参照モデル 223 の各状態  $I$  ( $I=1, 2, 3$ ) における出力分布、即ち、上記数 45 に示される出力分布に対する標準モデルの確率 (ここでは、上記数 51 に示される尤度  $\log L$ ) を極大化もしくは最大化するような標準モデルの統計量 (上記数 42 に示される混合重み係数、上記数 43 に示される平均値、および、上記数 44 に示される分散値) を推定する。

## 【0095】

ただし、本実施の形態では、上記数 45 に示された出力分布における

## 【数 6 5】

$$L_i \quad (i=1,2,\dots,N)$$

は、5 (各参照モデルの混合数) である。

## 【0096】

具体的には、上記数 52、数 53 および数 54 に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

このとき、統計量推定部 206 d の第 2 近似部 206 e により、上記数 55 に示される近似式が用いられる。

## 【0097】

ここで、第 2 近似部 206 e は、第 1 の実施の形態と異なり、上記数 55 の近似式の分母に示された出力分布

【数 6 6】

$$\omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2) \quad (m = 1, 2, \dots, M)$$

の中から、上記数 5 5 の近似式の分子に示された出力分布

【数 6 7】

$$\omega_{(k)} f(x; \mu_{(k)}, \sigma_{(k)}^2)$$

に距離的に近い 3 個 ( $P = 3$ ) の出力分布

【数 6 8】

$$\omega_{(p)} f(x; \mu_{(p)}, \sigma_{(p)}^2) \quad (p = 1, 2, \dots, P)$$

を選択し、選択した 3 個の出力分布を用いて、上記数 5 6 に示された単一ガウス分布の平均値 (数 5 7) および分散値 (数 5 8) を、それぞれ、以下の数 6 9 および数 7 0 に示された式に従って算出する。

【0 0 9 8】

【数 6 9】

$$\mu = \frac{\sum_{p=1}^P \omega_{(p)} \mu_{(p)}}{\sum_{p=1}^P \omega_{(p)}}$$

【0 0 9 9】

【数 7 0】

$$\sigma^2 = \frac{\sum_{p=1}^P \omega_{(p)} (\sigma_{(p)}^2 + \mu_{(p)}^2)}{\sum_{p=1}^P \omega_{(p)}} - \mu^2$$

図 9 は、第 2 近似部 2 0 6 e による近似計算を説明する図である。第 2 近似部 2 0 6 e は、本図に示されるように、上記数 5 5 に示された近似式における単一ガウス分布 (数 5 6) を、標準モデルを構成する M 個の混合ガウス分布の中から

、計算対象となる混合ガウス分布に近い一部（P 個）の混合ガウス分布だけを用いて決定している。したがって、全部（M 個）の混合ガウス分布を用いる第 1 の実施の形態と比較し、近似計算における計算量が削減される。

### 【0100】

以上の第 2 近似部 206 e による近似式を考慮してまとめると、統計量推定部 206 d での計算式は次の通りになる。つまり、統計量推定部 206 d は、以下の数 71、数 72 および数 73 に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、統計量記憶部 206 c に記憶する。そして、このような統計量の推定と統計量記憶部 206 c への記憶を R（ $\geq 1$ ）回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル 222 の統計量として出力する。なお、繰り返し計算においては、その回数に対応させて、上記近似計算における出力分布の選択個数 P を小さくし、最終的に  $P = 1$  とする計算を行う。

### 【0101】

#### 【数 71】

$$\omega_{(k)} = \frac{\left( \sum_{i=1}^N \sum_{l=1}^{L_i} \alpha_{(k,l,i)} \right)}{\sum_{k=1}^M \omega_{(k)} \left( \sum_{i=1}^N \sum_{l=1}^{L_i} \alpha_{(k,l,i)} \right)} \omega_{(k)}$$

$$\alpha_{(k,l,i)} = v_{(l,i)} \prod_{j=1}^J A_{(k,l,i,j)}$$

$$A_{(k,l,i,j)} = \frac{\sigma_{(j)}^2}{\sqrt{\sigma_{(k,j)}^2 \sigma_{(l,j)}^2 + \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 - \sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2}} \\ \times \exp \left\{ \frac{1}{2} \left( \frac{\left( \frac{\sigma_{(k,j)} \sigma_{(l,i,j)}}{\sigma_{(j)}} \mu_{(j)} - \frac{\sigma_{(k,j)} \sigma_{(j)}}{\sigma_{(l,i,j)}} \mu_{(l,i,j)} - \frac{\sigma_{(l,i,j)} \sigma_{(j)}}{\sigma_{(k,j)}} \mu_{(k,j)} \right)^2}{\sigma_{(k,j)}^2 \sigma_{(l,j)}^2 + \sigma_{(l,i,j)}^2 \sigma_{(j)}^2 - \sigma_{(k,j)}^2 \sigma_{(l,i,j)}^2} + \frac{\mu_{(j)}^2}{\sigma_{(j)}^2} - \frac{\mu_{(l,i,j)}^2}{\sigma_{(l,i,j)}^2} - \frac{\mu_{(k,j)}^2}{\sigma_{(k,j)}^2} \right) \right\}$$

### 【0102】

【数 7 2】

$$\mu_{(k,p)} = \frac{\sum_{i=1}^N \sum_{l=1}^{L_i} \beta_{(k,l,i,p)} \alpha_{(k,l,i)}}{\left( \sum_{i=1}^N \sum_{l=1}^{L_i} \alpha_{(k,l,i)} \right)}$$

$$\beta_{(k,l,i,p)} = \frac{\sigma_{(k,p)}^2 \sigma_{(p)}^2 \mu_{(l,i,p)} + \sigma_{(l,i,p)}^2 \sigma_{(p)}^2 \mu_{(k,p)} - \sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2 \mu_{(p)}}{\sigma_{(k,p)}^2 \sigma_{(p)}^2 + \sigma_{(l,i,p)}^2 \sigma_{(p)}^2 - \sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2}$$

【0103】

【数 7 3】

$$\sigma_{(k,p)}^2 = \frac{\sum_{i=1}^N \sum_{l=1}^{L_i} \gamma_{(k,l,i,p)} \alpha_{(k,l,i)}}{\left( \sum_{i=1}^N \sum_{l=1}^{L_i} \alpha_{(k,l,i)} \right)}$$

$$\gamma_{(k,l,i,p)} = \left\{ \frac{\sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2 \sigma_{(p)}^2}{\sigma_{(k,p)}^2 \sigma_{(p)}^2 + \sigma_{(l,i,p)}^2 \sigma_{(p)}^2 - \sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2} + \left( \mu_{(k,p)} - \frac{\sigma_{(k,p)}^2 \sigma_{(p)}^2 \mu_{(l,i,p)} + \sigma_{(l,i,p)}^2 \sigma_{(p)}^2 \mu_{(k,p)} - \sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2 \mu_{(p)}}{\sigma_{(k,p)}^2 \sigma_{(p)}^2 + \sigma_{(l,i,p)}^2 \sigma_{(p)}^2 - \sigma_{(k,p)}^2 \sigma_{(l,i,p)}^2} \right)^2 \right\}$$

なお、状態遷移確率については、HMMの対応する状態遷移確率を参照モデル 223 に対して全て加えあわせた全体が 1 になるように正規化したものを用いる。

【0104】

以上説明したように、本発明の第 2 の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

【0105】

なお、標準モデル 222 は、音素ごとに HMM を構成するに限らず、文脈依存の HMM で構成してもよい。

また、標準モデル 222 を構成する HMM は、音素ごとに異なる状態数により構成してもよいし、状態ごとに異なる混合数の混合ガウス分布により構成してもよい。

## 【0106】

また、参照モデル 221 は、話者ごと HMM において、異なる状態数により構成してもよいし、異なる混合数の混合ガウス分布により構成してもよい。

また、参照モデル 221 は、話者ごと HMM に限らず、話者・雑音・声の調子ごとに作成してもよい。

## 【0107】

また、標準モデル 222 を CD-ROM、ハードディスク、DVD-RAM などのストレージデバイスに記録してもよい。

また、参照モデル 221 を作成する代わりに、CD-ROM、DVD-RAM などのストレージデバイスから読み込んでもよい。

## 【0108】

また、参照モデル選択部 205 は、利用情報 224 に基づいて利用者ごとに選択する参照モデルの数を変えてもよい。

また、参照モデル準備部 202 は、必要に応じて新たな参照モデルを作成して参照モデル記憶部 203 に追加・更新してもよいし、参照モデル記憶部 203 に格納されている不要な参照モデルを削除してもよい。

## 【0109】

また、参照モデル準備部 202 は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部 203 に追加・更新してもよい。

また、上記近似計算において選択する出力分布の個数 P は、対象とする事象によって異なってもよいし、分布間距離に基づいて決定してもよい。

## 【0110】

(第 3 の実施の形態)

図 10 は、本発明の第 3 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置が PDA (Personal Digital Assistant) 301 に組み込まれた例が示されている。本実施の形態では雑音識別用の標準モデル (雑音モデル) を作成する場合を例にして説明する。

## 【0111】

PDA301は、携帯情報端末であり、事象の出力確率によって定義される雑音識別用の標準モデルを作成する標準モデル作成装置として、読み込み部311と、参照モデル準備部302と、参照モデル記憶部303と、利用情報作成部304と、参照モデル選択部305と、標準モデル作成部306と、仕様情報作成部307と、マイク312と、雑音識別部313とを備える。

#### 【0112】

読み込み部311は、CD-ROMなどのストレージデバイスに書き込まれた乗用車Aの参照モデル、乗用車Bの参照モデル、バスAの参照モデル、小雨の参照モデル、大雨の参照モデルなどの雑音の参照モデルを読み込む。参照モデル準備部302は、読み込まれた参照モデル321を参照モデル記憶部303へ送信する。参照モデル記憶部303は、参照モデル321を記憶する。

#### 【0113】

利用情報作成部304は、利用情報324である雑音の種類をPDA301の画面とキーを利用して作成する。参照モデル選択部305は、利用情報324である雑音の種類に音響的に近い参照モデルを、参照モデル記憶部303が記憶している参照モデル321の中から選択する。仕様情報作成部307は、PDA301の仕様に基づき仕様情報325を作成する。ここで、仕様情報とは、作成する標準モデルの仕様に関する情報であり、ここでは、PDA301が備えるCPUの処理能力に関する情報である。

#### 【0114】

標準モデル作成部306は、仕様情報作成部307で作成された仕様情報325に基づいて、参照モデル選択部305が選択した雑音の参照モデル323に対する確率または尤度を最大化または極大化するように標準モデル322を作成する処理部であり、標準モデルの構造（ガウス分布の混合数など）を決定する標準モデル構造決定部306aと、標準モデルを計算するための統計量の初期値を決定することで初期標準モデルを作成する初期標準モデル作成部306bと、決定された初期標準モデルを記憶する統計量記憶部306cと、統計量記憶部306cに記憶された初期標準モデルに対して、第3近似部306eによる近似計算等を用いることにより、参照モデル選択部305が選択した参照モデル323に対

する確率または尤度を最大化または極大化するような統計量を算出する（最終的な標準モデルを生成する）統計量推定部 306d とからなる。

#### 【0115】

雑音識別部 313 は、標準モデル作成部 306 で作成された標準モデル 322 を用いて、マイク 312 から入力された雑音の種類を識別する。

次に、以上のように構成された PDA 301 の動作について説明する。

#### 【0116】

図 11 は、PDA 301 の動作手順を示すフローチャートである。

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（ステップ S300）。つまり、読み込み部 311 は、ストレージデバイスに書き込まれた雑音の参照モデルを読み込み、参照モデル準備部 302 は、読み込まれた参照モデル 321 を参照モデル記憶部 303 へ送信し、参照モデル記憶部 303 は、参照モデル 321 を記憶する。

#### 【0117】

参照モデル 321 は、GMM より構成される。参照モデル 321 の一例を図 12 に示す。ここでは、各雑音モデルは混合数が 3 個の GMM により構成される。特徴量として 5 次元 ( $J=5$ ) の LPC ケプストラム係数が用いられる。

#### 【0118】

次に、利用情報作成部 304 は、識別したい雑音の種類である利用情報 324 を作成する（ステップ S301）。図 13 に PDA 301 の選択画面の一例を示す。ここでは、乗用車の雑音が選択される。参照モデル選択部 305 は、選択された利用情報 324 である乗用車の雑音に音響的に近い参照モデルである乗用車 A の参照モデルと乗用車 B の参照モデルを、参照モデル記憶部 303 が記憶している参照モデル 321の中から選択する（ステップ S302）。

#### 【0119】

そして、仕様情報作成部 307 は、PDA 301 の仕様に基づき、仕様情報 325 を作成する（ステップ S303）。ここでは、PDA 301 の CPU の仕様に基づき CPU パワーが小さいという仕様情報 325 を作成する。標準モデル作成部 306 は、作成された仕様情報 325 に基づいて、参照モデル選択部 305

が選択した参照モデル 323 に対する確率または尤度を最大化または極大化するように標準モデル 322 を作成する (ステップ S304)。

#### 【0120】

最後に、雑音識別部 313 は、利用者によってマイク 312 から入力された雑音に対して、標準モデル 322 を用いて、雑音の識別を行う (ステップ S305)。

#### 【0121】

次に、図 11 におけるステップ S304 (標準モデルの作成) の詳細な手順を説明する。手順の流れは、図 4 に示されたフローチャートと同様である。ただし、採用する標準モデルの構造や具体的な近似計算等が異なる。

#### 【0122】

まず、標準モデル構造決定部 306a は、標準モデルの構造を決定する (図 4 のステップ S102)。ここでは、標準モデルの構造として、仕様情報 325 である CPU パワーが小さいという情報に基づいて 1 混合 ( $M=1$ ) の GMM により標準モデル 322 を構成すると決定する。

#### 【0123】

次に、初期標準モデル作成部 306b は、標準モデルを計算するための統計量の初期値を決定する (図 4 のステップ S102b)。ここでは、選択された参照モデル 323 である乗用車 A の 3 混合の参照モデルを、統計処理計算を用いて 1 つのガウス分布に統合したものを統計量の初期値として統計量記憶部 306c に記憶する。

#### 【0124】

具体的には、初期標準モデル作成部 306b は、上記数 39 に示される出力分布を生成する。

ただし、本実施の形態では、上記数 39 に示された出力分布における

#### 【0125】

#### 【数 74】

$x$

は、5 次元 ( $J=5$ ) の LPC ケプストラム係数を表す。



## 【0126】

そして、統計量推定部306dは、参照モデル選択部305が選択した2個の参照モデル323を用いて、統計量記憶部306cに記憶された標準モデルの統計量を推定する（図4のステップS102c）。

## 【0127】

つまり、2個（ $N=2$ ）の参照モデル323における出力分布、即ち、上記数45に示される出力分布に対する標準モデルの確率（ここでは、上記数51に示される尤度 $\log L$ ）を極大化もしくは最大化するような標準モデルの統計量（上記数42に示される混合重み係数、上記数43に示される平均値、および、上記数44に示される分散値）を推定する。

## 【0128】

ただし、本実施の形態では、上記数45に示された出力分布における  
【数75】

$$L_i \quad (i=1,2,\dots,N)$$

は、3（各参照モデルの混合数）である。

## 【0129】

具体的には、上記数52、数53および数54に示される式に従って、それぞれ、標準モデルの混合重み係数、平均値および分散値を算出する。

このとき、統計量推定部306dの第3近似部306eは、標準モデルの各ガウス分布はお互いにはほとんど影響を与えないと仮定して、以下の近似式を用いる。

## 【0130】

【数76】

$$\gamma(x,k) = \frac{\omega_{(k)} f(x; \mu_{(k)}, \sigma_{(k)}^2)}{\sum_{m=1}^M \omega_{(m)} f(x; \mu_{(m)}, \sigma_{(m)}^2)} \approx 1.0$$

また、参照モデル323のガウス分布

【数 7 7】

$$g_i(x; \mu_{(i)}, \sigma_{(i)}^2) \quad (i=1, 2, \dots, N)$$

は、マハラノビス距離、カルバック・ライブラー (KL) 距離などの分布間距離が最も近い標準モデルのガウス分布

【数 7 8】

$$f(x; \mu_{(m)}, \sigma_{(m)}^2)$$

にのみ影響すると仮定する。

【0 1 3 1】

図 1 4 は、この統計量推定部 3 0 6 d による統計量の推定手順を示す概念図である。各参照モデルの各ガウス分布に対して、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 k であるガウス分布を用いて統計量の推定を行うことが示されている。

【0 1 3 2】

図 1 5 は、第 3 近似部 3 0 6 e による近似計算を説明する図である。第 3 近似部 3 0 6 e は、本図に示されるように、各参照モデルの各ガウス分布に対して、距離が最も近い標準モデルのガウス分布 k を決定することで、上記数 7 6 に示された近似式を用いている。

【0 1 3 3】

以上の第 3 近似部 3 0 6 e による近似式を考慮してまとめると、統計量推定部 3 0 6 d での計算式は次の通りになる。つまり、統計量推定部 3 0 6 d は、以下の数 7 9、数 8 0 および数 8 1 に示される式に従って、それぞれ、混合重み係数、平均値および分散値を算出し、それらのパラメータによって特定される標準モデルを最終的な標準モデル 3 2 2 として生成する。

【0 1 3 4】

【数 79】

$$\omega_{(k)} = \frac{\sum_{i^{\min}} \sum_{l^{\min}} v_{(l^{\min}, i^{\min})}}{\sum_i^N \sum_l^{L_i} v_{(l, i)}}$$

$$(k = 1, 2, \dots, M)$$

(ここで、分子の和は、各参照モデルの各ガウス分布に対して、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布  $k$  であるガウス分布に関する和を意味する。)

【0135】

【数 80】

$$\mu_{(k, j)} = \frac{\sum_{i^{\min}} \sum_{l^{\min}} v_{(l^{\min}, i^{\min}, j)} \mu_{(l^{\min}, i^{\min}, j)}}{\sum_{i^{\min}} \sum_{l^{\min}} v_{(l^{\min}, i^{\min}, j)}}$$

$$(k = 1, 2, \dots, M, j = 1, 2, \dots, J)$$

(ここで、分母、分子の和は、各参照モデルの各ガウス分布に対して、マハラノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布  $k$  であるガウス分布に関する和を意味する。)

【0136】

【数 81】

$$\sigma_{(k, j)}^2 = \frac{\sum_{i^{\min}} \sum_{l^{\min}} v_{(l^{\min}, i^{\min}, j)} (\sigma_{(l^{\min}, i^{\min}, j)}^2 + \mu_{(l^{\min}, i^{\min}, j)}^2)}{\sum_{i^{\min}} \sum_{l^{\min}} v_{(l^{\min}, i^{\min}, j)}} - \mu_{(k, j)}^2$$

$$(k = 1, 2, \dots, M, j = 1, 2, \dots, J)$$

(ここで、分母、分子の和は、各参照モデルの各ガウス分布に対して、マハラ

ノビス距離などの分布間距離が最も近いものが標準モデルのガウス分布 $k$ であるガウス分布に関する和を意味する。)

統計量推定部 306 d は、このように推定した標準モデルの統計量を統計量記憶部 306 c に記憶する。そして、このような統計量の推定と統計量記憶部 306 c への記憶を  $R$  ( $\geq 1$ ) 回、繰り返す。その結果得られた統計量を最終的に生成する標準モデル 322 の統計量として出力する。

#### 【0137】

以上説明したように、本発明の第3の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

#### 【0138】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、統計量推定部 306 d による処理の繰り返し回数は、上記数 51 に示された尤度の大きさがある一定のしきい値以上になるまでの回数としてもよい。

#### 【0139】

また、標準モデル 322 を構成する GMM は、雑音の種類ごとに異なる混合数の混合ガウス分布により構成してもよい。

また、識別モデルは、雑音モデルに限らず、話者を識別してもよいし、年齢などを識別してもよい。

#### 【0140】

また、標準モデル 322 を CD-ROM、DVD-RAM、ハードディスクなどのストレージデバイスに記録してもよい。

また、参照モデル 321 を CD-ROM などのストレージデバイスから読み込む代わりに、PDA 301 において雑音データから参照モデル 321 を作成してもよい。

#### 【0141】

また、参照モデル準備部 302 は、必要に応じて CD-ROM などのストレージ

ジデバイスから読み込まれた新たな参照モデルを参照モデル記憶部 303 に追加・更新してもよいし、参照モデル記憶部 303 に格納されている不要な参照モデルを削除してもよい。

#### 【0142】

また、参照モデル準備部 302 は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部 303 に追加・更新してもよい。

#### 【0143】

(第 4 の実施の形態)

図 16 は、本発明の第 4 の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ 401 に組み込まれた例が示されている。本実施の形態では顔認識用の標準モデルを作成する場合を例にして説明する。

#### 【0144】

サーバ 401 は、通信システムにおけるコンピュータ装置等であり、事象の出力確率によって定義される顔認識用の標準モデルを作成する標準モデル作成装置として、カメラ 411 と、画像データ蓄積部 412 と、参照モデル準備部 402 と、参照モデル記憶部 403 と、利用情報受信部 404 と、参照モデル選択部 405 と、標準モデル作成部 406 と、書き込み部 413 とを備える。

#### 【0145】

カメラ 411 により、顔の画像データが収集され、画像データ蓄積部 412 に顔画像データが蓄積される。参照モデル準備部 402 は、画像データ蓄積部 412 が蓄積した顔画像データを用いて話者ごとに参照モデル 421 を作成し、参照モデル記憶部 403 に記憶する。

#### 【0146】

利用情報受信部 404 は、利用者が希望する顔認識の対象となる人間の年齢の年代と性別の情報を利用情報 424 として電話 414 により受信する。参照モデル選択部 405 は、利用情報受信部 404 が受信した利用情報 424 に基づいて、参照モデル記憶部 403 が記憶している参照モデル 421 の中から、利用情報 424 が示す年代と性別の話者に対応する参照モデル 423 を選択する。

## 【0147】

標準モデル作成部406は、参照モデル選択部405が選択した話者の顔画像の参照モデル423に対する確率または尤度を最大化または極大化するように標準モデル422を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有するとともに、第1の実施の形態における第1近似部104eと第3の実施の形態における第3近似部306eの機能を有する。つまり、第1～第3の実施の形態で示された3種類の近似計算を組み合わせた計算を行う。

## 【0148】

書き込み部413は、標準モデル作成部406が作成した標準モデル422をCD-ROMなどのストレージデバイスに書き込む。

次に、以上のように構成されたサーバ401の動作について説明する。

## 【0149】

図17は、サーバ401の動作手順を示すフローチャートである。図18は、サーバ401の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

## 【0150】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図17のステップS400）。つまり、カメラ411によりAさんからZさんの顔画像データを収集して画像データ蓄積部412に蓄積する。参照モデル準備部402は、画像データ蓄積部412が蓄積した顔画像データを用いて、話者ごとの参照モデル421をEMアルゴリズムにより作成する。ここでは参照モデル421はGMMで構成される。

## 【0151】

参照モデル記憶部403は、参照モデル準備部402が作成した参照モデル421を記憶する。ここでは、図18の参照モデル421に示されるように、AさんからZさんの全ての参照モデルが、混合数が5個のGMMにより構成される。特徴量として100次元（ $J=100$ ）の画素の濃度値を用いる。

## 【0152】

次に、利用情報受信部 404 は、利用情報 424 である年代と性別の情報を電話 414 により受信する（図 17 のステップ S401）。ここでは、利用情報 424 として、11 歳から 15 歳の男性と 22 歳から 26 歳の女性である。参照モデル選択部 405 は、その利用情報 424 に基づいて、参照モデル記憶部 403 が記憶している参照モデル 421 から、利用情報 424 に対応する参照モデル 423 を選択する（図 17 のステップ S402）。具体的には、図 18 の「選択された参照モデル 423」に示されるように、ここでは、11 歳から 15 歳の男性および 22 歳から 26 歳の女性の参照モデルを選択する。

#### 【0153】

そして、標準モデル作成部 406 は、参照モデル選択部 405 が選択した話者の参照モデル 423 に対する確率または尤度を最大化または極大化するように標準モデル 422 を作成する（図 17 のステップ S403）。ここでは、図 18 の標準モデル 422 に示されるように、2 つの標準モデル 422 それぞれを、混合数が 3 個の GMM により構成する。

#### 【0154】

標準モデル 422 の作成方法は、基本的には、第 2 の実施の形態と同様に行われる。ただし、標準モデル 422 の統計量の推定における近似計算については、具体体には、以下のようにして行われる。つまり、標準モデル作成部 406 は、内蔵の記憶部等を介することで、第 1 の実施の形態における第 1 近似部 104e による近似計算と同様の近似計算によって作成したモデルを初期値として、第 2 の実施の形態における第 2 近似部 206e による近似計算と同様の近似計算による計算を行い、その結果を初期値として第 3 の実施の形態における第 3 近似部 306e による近似計算と同様の近似計算を行う。

#### 【0155】

書き込み部 413 は、標準モデル作成部 406 が作成した 2 つの標準モデル 422 を CD-ROM などのストレージデバイスに書き込む（図 17 のステップ S404）。

#### 【0156】

利用者は、11 歳から 15 歳の男性の標準モデルと 22 歳から 26 歳の女性の

標準モデルが書き込まれたストレージデバイスを郵送で受け取る。

以上説明したように、本発明の第4の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

#### 【0157】

なお、標準モデル422を構成するGMMは、話者ごとに異なる混合数の混合ガウス分布により構成してもよい。

また、参照モデル準備部402は、必要に応じて新たな参照モデルを作成して参照モデル記憶部403に追加・更新してもよいし、参照モデル記憶部403に格納されている不要な参照モデルを削除してもよい。

#### 【0158】

(第5の実施の形態)

図19は、本発明の第5の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ501に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

#### 【0159】

サーバ501は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部511と、音声データ蓄積部512と、参照モデル準備部502と、参照モデル記憶部503と、利用情報受信部504と、参照モデル選択部505と、標準モデル作成部506と、仕様情報受信部507と、書き込み部513とを備える。

#### 【0160】

読み込み部511は、CD-ROMなどのストレージデバイスに書き込まれた子供、成人、高齢者の音声データを読み込み、音声データ蓄積部512に蓄積する。参照モデル準備部502は、音声データ蓄積部512が蓄積した音声データ



を用いて話者ごとに参照モデル 521 を作成する。参照モデル記憶部 503 は、参照モデル準備部 502 が作成した参照モデル 521 を記憶する。

#### 【0161】

仕様情報受信部 507 は、仕様情報 525 を受信する。利用情報受信部 504 は、利用情報 524 である利用者の音声を受信する。参照モデル選択部 505 は、利用情報 524 である利用者の音声に音響的に近い話者の参照モデルを、参照モデル記憶部 503 が記憶している参照モデル 521 から選択する。

#### 【0162】

標準モデル作成部 506 は、仕様情報 525 に基づいて、参照モデル選択部 505 が選択した話者の参照モデル 523 に対する確率または尤度を最大化または極大化するように標準モデル 522 を作成する処理部であり、第 1 の実施の形態における標準モデル作成部 104 と同一の機能を有する。書き込み部 513 は、標準モデル作成部 506 が作成した標準モデル 522 を CD-ROM などのストレージデバイスに書き込む。

#### 【0163】

次に、以上のように構成されたサーバ 501 の動作について説明する。

図 20 は、サーバ 501 の動作手順を示すフローチャートである。図 21 は、サーバ 501 の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

#### 【0164】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図 20 のステップ S500）。つまり、読み込み部 511 は、CD-ROM などのストレージデバイスに書き込まれた音声データを読み込み、音声データ蓄積部 512 に蓄積する。参照モデル準備部 502 は、音声データ蓄積部 512 が蓄積した音声データを用いて話者ごとに参照モデル 521 をバウム・ウェルチの再推定の方法により作成する。参照モデル記憶部 503 は、参照モデル準備部 502 が作成した参照モデル 521 を記憶する。

#### 【0165】

参照モデル 521 は、音素ごとの HMM により構成される。ここでは、図 21

の参照モデル 521 に示されるように、子供の各話者の参照モデルは、状態数 3 個、各状態は混合数が 3 個の混合ガウス分布により HMM の出力分布が構成され、成人の各話者の参照モデルが、状態数 3 個、各状態は混合数が 64 個の混合ガウス分布により HMM の出力分布が構成され、高齢者の各話者の参照モデルは、状態数 3 個、各状態は混合数が 16 個の混合ガウス分布により HMM の出力分布が構成される。これは、子供の音声データが比較的少なく、成人の音声データが多いためである。特徴量として 25 次元 ( $J=25$ ) のメルケプストラム係数が用いられる。

#### 【0166】

次に、利用情報受信部 504 は、利用者の音声を、端末装置 514 から、利用情報 524 として受信する (図 20 のステップ S501)。参照モデル選択部 505 は、利用情報 524 である利用者の音声に音響的に近い参照モデル 523 を、参照モデル記憶部 503 が記憶している参照モデル 521 から選択する (図 20 のステップ S502)。具体的には、図 21 の「選択された参照モデル 523」に示されるように、ここでは、近い話者 10 人 ( $N=10$ ) の参照モデルが選択される

そして、仕様情報受信部 507 は、利用者の要求に基づき仕様情報 525 を端末装置 514 から受信する (図 20 のステップ S503)。ここでは、速い認識処理という仕様情報 525 を受信する。標準モデル作成部 506 は、仕様情報受信部 507 が受信した仕様情報 525 に基づいて、参照モデル選択部 505 が選択した話者の参照モデル 523 に対する確率または尤度を最大化または極大化するように標準モデル 522 を作成する (図 20 のステップ S504)。具体的には、標準モデル 522 は、図 21 の標準モデル 522 に示されるように、仕様情報 525 である速い認識処理という情報に基づいて、2 混合 ( $M=2$ ) で、3 状態の HMM より構成する。HMM は音素ごとに構成する。

#### 【0167】

標準モデル 522 の作成方法は、第 1 の実施の形態と同様に行われる。

書き込み部 513 は、標準モデル作成部 506 が作成した標準モデル 522 を CD-ROM などのストレージデバイスに書き込む (図 20 のステップ S505)

)。

#### 【0168】

以上説明したように、本発明の第5の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

#### 【0169】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

また、参照モデル準備部502において、参照モデルごとにデータ数に適した混合数の高精度な参照モデルを準備でき、高精度な参照モデルを用いて標準モデルを作成できる。このため高精度な標準モデルの利用が可能となる。

#### 【0170】

なお、標準モデル522は、音素ごとにHMMを構成するに限らず、文脈依存のHMMで構成してもよい。

また、標準モデル522を構成するHMMは、状態ごとに異なる混合数の混合ガウス分布により構成してもよい。

#### 【0171】

また、標準モデル522を用いて、サーバ501において音声認識を行ってもよい。

また、参照モデル準備部502は、必要に応じて新たな参照モデルを作成して参照モデル記憶部503に追加・更新してもよいし、参照モデル記憶部503に格納されている不要な参照モデルを削除してもよい。

#### 【0172】

##### (第6の実施の形態)

図22は、本発明の第6の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ601に組み込まれた例が示されている。本実施の形態では意図理解のための標準モデル（嗜好モデル）を作成する場合を例にし

て説明する。

【0173】

サーバ601は、通信システムにおけるコンピュータ装置等であり、事象の出力確率によって定義される意図理解用の標準モデルを作成する標準モデル作成装置として、読み込み部611と、参照モデル準備部602と、参照モデル記憶部603と、利用情報受信部604と、参照モデル選択部605と、標準モデル作成部606と、仕様情報作成部607とを備える。

【0174】

読み込み部611は、CD-ROMなどのストレージデバイスに書き込まれた年齢別の話者Aさんから話者Zさんの嗜好モデルを読み込み、参照モデル準備部602は、読み込まれた参照モデル621を参照モデル記憶部603へ送信し、参照モデル記憶部603は、参照モデル621を記憶する。

【0175】

仕様情報作成部607は、普及しているコンピュータのCPUパワーに合わせて仕様情報625を作成する。利用情報受信部604は、端末装置614から利用情報624を受信する。参照モデル選択部605は、利用情報受信部604が受信した利用情報624に基づいて、参照モデル記憶部603が記憶している参照モデル621からの中から、利用情報624に対応した参照モデル623を選択する。

【0176】

標準モデル作成部606は、仕様情報作成部607が作成した仕様情報625に基づいて、参照モデル選択部605が選択した参照モデル623に対する確率または尤度を最大化または極大化するように標準モデル622を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有するとともに、第3の実施の形態における第3近似部306eの機能を有する。つまり、第1および第3の実施の形態で示された2種類の近似計算を組み合わせた計算を行う。

【0177】

次に、以上のように構成されたサーバ601の動作について説明する。

図 23 は、サーバ 601 の動作手順を示すフローチャートである。図 24 は、サーバ 601 の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

#### 【0178】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図 23 のステップ S600）。つまり、読み込み部 611 は、CD-ROM などのストレージデバイスに書き込まれた年齢別の話者 Aさんから話者 Zさんの嗜好モデルを読み込み、参照モデル準備部 602 は、読み込まれた参照モデル 621 を参照モデル記憶部 603 へ送信し、参照モデル記憶部 603 は、参照モデル 621 を記憶する。

#### 【0179】

参照モデル 621 は、GMM より構成される。ここでは、図 24 の参照モデル 621 に示されるように、混合数が 3 個の GMM により構成される。入力データとして、趣味、性格などを数値化した 5 次元 ( $J=5$ ) の特徴量を用いる。

#### 【0180】

次に、利用情報受信部 604 は、嗜好モデルを作成したい年齢層である利用情報 624 を受信する（図 23 のステップ S601）。ここでは、20代、30代、40代の年代別の嗜好モデルを利用するという利用情報 624 である。参照モデル選択部 605 は、図 24 の「選択された参照モデル 623」に示されるように、利用情報受信部 604 が受信した利用情報 624 が示す年代の話者の嗜好モデルを、参照モデル記憶部 603 が記憶している参照モデル 621 から選択する（図 23 のステップ S602）。

#### 【0181】

そして、仕様情報作成部 607 は、普及しているコンピュータの CPU パワー、記憶容量などに基づき仕様情報 625 を作成する（図 23 のステップ S603）。ここでは、通常速度の認識処理という仕様情報 625 を作成する。

#### 【0182】

標準モデル作成部 606 は、仕様情報作成部 607 が作成した仕様情報 625 に基づいて、参照モデル選択部 605 が選択した話者の参照モデル 623 に対す

る確率または尤度を最大化または極大化するように標準モデル 622 を作成する (図 23 のステップ S604)。ここでは、標準モデル 622 は、図 24 の標準モデル 622 に示されるように、仕様情報 625 である通常速度の認識処理という情報に基づいて 3 混合 ( $M=3$ ) の GMM より構成する。

#### 【0183】

標準モデル 622 の作成方法は、基本的には、第 2 の実施の形態と同様に行われる。ただし、標準モデル 622 の統計量の推定における近似計算については、具体体には、以下のようにして行われる。つまり、標準モデル作成部 606 は、内蔵の記憶部等を介することで、第 2 の実施の形態における第 2 近似部 206 e による近似計算と同様の近似計算による計算を行い、その結果を初期値として第 3 の実施の形態における第 3 近似部 306 e による近似計算と同様の近似計算を行う。

#### 【0184】

以上説明したように、本発明の第 6 の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

#### 【0185】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、標準モデル 622 を構成する GMM は、話者ごとに異なる混合数の混合ガウス分布により構成してもよい。

#### 【0186】

また、参照モデル準備部 602 は、必要に応じて CD-ROM などのストレージデバイスから読み込まれた新たな参照モデルを参照モデル記憶部 603 に追加・更新してもよいし、参照モデル記憶部 603 に格納されている不要な参照モデルを削除してもよい。

#### 【0187】

また、参照モデルおよび標準モデルの GMM はベイジアンネットの一部を表現

するものでもよい。

#### 【0188】

(第7の実施の形態)

図25は、本発明の第7の実施の形態における標準モデル作成装置の全体構成を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ701に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

#### 【0189】

サーバ701は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部711と、参照モデル準備部702と、参照モデル記憶部703と、利用情報受信部704と、参照モデル選択部705と、標準モデル作成部706と、仕様情報受信部707と、標準モデル記憶部708と、標準モデル送信部709とを備える。

#### 【0190】

参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMなどのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する。

#### 【0191】

仕様情報受信部707は、端末装置712から仕様情報725を受信する。利用情報受信部704は、端末装置712から利用情報724である雑音下で発声した利用者の音声を受信する。参照モデル選択部705は、利用情報724である利用者の音声に音響的に近い話者・雑音・声調子の参照モデル723を、参照モデル記憶部703が記憶している参照モデル721の中から選択する。

#### 【0192】

標準モデル作成部706は、仕様情報受信部707が受信した仕様情報725に基づいて、参照モデル選択部705が選択した参照モデル723に対する確率

または尤度を最大化または極大化するように標準モデル722を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有する。標準モデル記憶部708は、仕様情報725に基づいた1もしくは複数の標準モデルを記憶する。標準モデル送信部709は、利用者の端末装置712から仕様情報と標準モデルの要求信号を受信すると、その仕様情報に適した標準モデルを端末装置712へ送信する。

#### 【0193】

次に、以上のように構成されたサーバ701の動作について説明する。

図26は、サーバ701の動作手順を示すフローチャートである。図27は、サーバ701の動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

#### 【0194】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図26のステップS700）。つまり、参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMなどのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する。ここでは、参照モデル721は、話者・雑音・声の調子ごとに、音素ごとのHMMにより構成される。また、各参照モデルは、図27の参照モデル721に示されるように、状態数3個、各状態は混合数が128個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として25次元（ $J=25$ ）のメルケプストラム係数が用いられる。

#### 【0195】

次に、利用情報受信部704は、利用者Aの雑音下での音声を端末装置712から利用情報724として受信する（図26のステップS701）。参照モデル選択部705は、利用情報724である利用者Aの音声に音響的に近い参照モデル723を、参照モデル記憶部703が記憶している参照モデル521の中から選択する（図26のステップS702）。具体的には、図27の「選択された参照モデル723」に示されるように、ここでは、近い話者100人（ $N=100$



) の参照モデルが選択される

そして、仕様情報受信部 707 は、利用者 A の要求に基づき仕様情報 725 を端末装置 712 から受信する (図 26 のステップ S703)。ここでは、高い認識精度という仕様情報 725 を受信する。標準モデル作成部 706 は、仕様情報 725 に基づいて、参照モデル選択部 705 が選択した参照モデル 723 に対する確率または尤度を最大化または極大化するように標準モデル 722 を作成する (図 26 のステップ S704)。具体的には、標準モデル 722 は、図 27 の標準モデル 722 に示されるように、仕様情報 725 である高い認識精度という情報に基づいて、64 混合 ( $M=64$ ) で、3 状態の HMM より構成する。HMM は音素ごとに構成する。

#### 【0196】

標準モデル 722 の作成方法は、第 2 の実施の形態と同様に行われる。

標準モデル記憶部 708 は、仕様情報 725 に基づいた 1 もしくは複数の標準モデル 722 を記憶する。ここでは、以前に作成した標準モデルである利用者 B の 16 混合の HMM がすでに記憶されており、新たに利用者 A の 64 混合の HMM が記憶される。

#### 【0197】

利用者 A は、端末装置 712 からサーバ 701 の標準モデル送信部 709 へ、仕様情報である利用者 A と雑音の種類と標準モデルの要求信号とを送信する (図 26 のステップ S706)。標準モデル送信部 709 は、利用者 A が送信した仕様情報と標準モデルの要求信号とを受信すると、その端末装置 712 へ、仕様に適した標準モデルを端末装置 712 へ送信する (図 26 のステップ S707)。ここでは、先ほど作成した利用者 A の標準モデル 722 を端末装置 712 へ送信する。

#### 【0198】

利用者 A は端末装置 712 において受信した標準モデル 722 を用いて音声認識を行う (図 26 のステップ S708)。

以上説明したように、本発明の第 7 の実施の形態によれば、利用情報に基づいて選択された複数の参照モデルに対する確率または尤度を最大化または極大化する

るように標準モデルの統計量を計算して標準モデルが作成されるので、利用状況によりふさわしい高精度な標準モデルが提供される。

#### 【0199】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

また、標準モデル記憶部708は、複数の標準モデルを記憶することができるため、必要に応じてすぐに標準モデルが提供される。

#### 【0200】

また、標準モデル送信部709により、標準モデルが端末装置712へ送信されるので、端末装置712とサーバ701が空間的に離れた場所に設置してある場合に、端末装置712は、容易にサーバ701が作成した標準モデルを利用することができる。

#### 【0201】

なお、標準モデル722は、音素ごとにHMMを構成するに限らず、文脈依存のHMMで構成してもよい。

また、標準モデル722を構成するHMMは、状態ごとに異なる混合数の混合ガウス分布により構成してもよい。

#### 【0202】

また、標準モデル722を用いて、サーバ701において音声認識を行い、認識結果を端末装置712へ送信してもよい。

また、参照モデル準備部702は、必要に応じて新たな参照モデルを作成して参照モデル記憶部703に追加・更新してもよいし、参照モデル記憶部703に格納されている不要な参照モデルを削除してもよい。

#### 【0203】

また、参照モデル準備部702は、必要に応じて、通信路を介して新たな参照モデルを参照モデル記憶部703に追加・更新してもよい。

#### 【0204】

(第8の実施の形態)

図28は、本発明の第8の実施の形態における標準モデル作成装置の全体構成

を示すブロック図である。ここでは、本発明に係る標準モデル作成装置がコンピュータシステムにおけるサーバ 801 に組み込まれた例が示されている。本実施の形態では音声認識用の標準モデル（適応モデル）を作成する場合を例にして説明する。

#### 【0205】

サーバ 801 は、通信システムにおけるコンピュータ装置等であり、事象の集合と事象または事象間の遷移の出力確率とによって定義される音声認識用の標準モデルを作成する標準モデル作成装置として、読み込み部 711 と、参照モデル準備部 702 と、参照モデル記憶部 703 と、利用情報受信部 704 と、参照モデル選択部 705 と、標準モデル作成部 706 と、仕様情報受信部 707 と、標準モデル記憶部 708 と、標準モデル送信部 709 と、参照モデル受信部 810 とを備える。

#### 【0206】

参照モデル準備部 702 は、読み込み部 711 が読み込んだ、CD-ROM などのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部 703 へ送信する。参照モデル記憶部 703 は、送信された参照モデル 721 を記憶する。また、参照モデル準備部 702 は、端末装置 712 からの送信に対して参照モデル受信部 810 が受信した音声認識用参照モデルを参照モデル記憶部 703 へ送信する。参照モデル記憶部 703 は、送信された参照モデル 721 を記憶する。

#### 【0207】

仕様情報受信部 707 は、端末装置 712 から仕様情報 725 を受信する。利用情報受信部 704 は、端末装置 712 から利用情報 724 である雑音下で発声した利用者の音声を受信する。参照モデル選択部 705 は、利用情報受信部 704 が受信した利用情報 724 である利用者の音声に音響的に近い話者・雑音・声調子の参照モデル 723 を、参照モデル記憶部 703 が記憶している参照モデル 721 から選択する。

#### 【0208】

標準モデル作成部 706 は、仕様情報 725 に基づいて、参照モデル選択部 7

05が選択した参照モデル723に対する確率または尤度を最大化または極大化するように標準モデル722を作成する処理部であり、第2の実施の形態における標準モデル作成部206と同一の機能を有する。標準モデル記憶部708は、仕様情報725に基づいた1もしくは複数の標準モデルを記憶する。標準モデル送信部709は、利用者の端末装置712から、仕様情報と標準モデルの要求信号とを受信すると、その端末装置712へ、仕様に適した標準モデルを送信する。

#### 【0209】

次に、以上のように構成されたサーバ801の動作について説明する。

図29は、サーバ801の動作手順を示すフローチャートである。なお、このサーバ801の動作手順を説明するための参照モデルおよび標準モデルの一例は、第7に実施の形態における図27と同様である。

#### 【0210】

まず、標準モデルの作成に先立ち、その基準となる参照モデルを準備する（図29のステップS800、S801）。つまり、参照モデル準備部702は、読み込み部711が読み込んだ、CD-ROMなどのストレージデバイスに書き込まれた話者・雑音・声の調子別の音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する（図29のステップS800）。ここでは、参照モデル721は、話者・雑音・声の調子ごとに、音素ごとのHMMにより構成される。また、参照モデル準備部702は、端末装置712が送信して参照モデル受信部810が受信した、利用者と端末装置712に適した音声認識用参照モデルを参照モデル記憶部703へ送信し、参照モデル記憶部703は、送信された参照モデル721を記憶する（図29のステップS801）。ここでは、各参照モデルは、図27の参照モデル721に示されるように、状態数3個、各状態は混合数が128個の混合ガウス分布によりHMMの出力分布が構成される。特徴量として25次元（ $J=25$ ）のメルケプストラム係数が用いられる。

#### 【0211】

以下、これらの参照モデル721を用いた標準モデル722の作成および端末

装置 712 への送信 (図 29 のステップ S802 ~ S809) は、第 7 の実施の形態における手順 (図 26 のステップ S701 ~ S708) と同様である。

#### 【0212】

図 30 は、本実施の形態における標準モデル作成装置を具体的に適用したシステム例を示す図である。ここには、インターネットや無線通信等を介して通信し合うサーバ 701 と端末装置 712 (携帯電話機 712a、カーナビゲーション装置 712b) とが示されている。

#### 【0213】

たとえば、携帯電話機 712a は、利用者の音声を利用情報とし、携帯電話機での利用である旨 (CPU の処理能力が低いこと) を仕様情報とし、予め記憶しているサンプルモデルを参照モデルとし、それら利用情報、仕様情報および参照モデルをサーバ 701 に送信することで、標準モデルの作成を要求する。その要求に対してサーバ 701 で標準モデルが作成されると、携帯電話機 712a は、その標準モデルをダウンロードし、その標準モデルを用いて利用者の音声を認識する。例えば、利用者の音声で、内部に保持するアドレス帳の名前と一致した場合には、その名前に対応する電話番号に自動発呼する。

#### 【0214】

また、カーナビゲーション装置 712b は、利用者の音声を利用情報とし、カーナビゲーション装置での利用である旨 (CPU の処理能力が通常であること) を仕様情報とし、予め記憶しているサンプルモデルを参照モデルとし、それら利用情報、仕様情報および参照モデルをサーバ 701 に送信することで、標準モデルの作成を要求する。その要求に対してサーバ 701 で標準モデルが作成されると、カーナビゲーション装置 712b は、その標準モデルをダウンロードし、その標準モデルを用いて利用者の音声を認識する。例えば、利用者の音声で、内部に保持する地名と一致した場合には、その地名を目標点とする現地点からの道順を示す地図を画面に自動表示する。

#### 【0215】

このようにして、携帯電話機 712a およびカーナビゲーション装置 712b は、自装置に適した標準モデルの作成をサーバ 701 に依頼することで、標準モ

デルの作成に必要な回路や処理プログラムを自装置内に実装する必要がなくなるとともに、様々な認識・認証対象の標準モデルを必要なタイミングで獲得することができる。

#### 【0216】

以上説明したように、本発明の第8の実施の形態によれば、参照モデル受信部810が受信した参照モデルを利用して標準モデルを作成できるため、高精度な標準モデルが提供される。

#### 【0217】

また、仕様情報に基づいて標準モデルが作成されるため、標準モデルを利用する機器にふさわしい標準モデルが準備される。

なお、参照モデル受信部810は、端末装置712とは異なる他の端末装置から参照モデルを受信してもよい。

#### 【0218】

また、第1～第8の実施の形態における標準モデルの統計量の近似計算については、各実施の形態における近似計算だけに限られず、第1～第3の実施の形態における合計3種類の近似計算の少なくとも1つを用いてもよい。つまり、3種類の近似計算のいずれであってもよいし、2以上の種類の近似計算の組み合わせであってもよい。

#### 【0219】

また、図30に示された応用例は、本実施の形態に限られるものではなく、他の実施の形態にも適用することができる。つまり、第1～第7の実施の形態で作成された標準モデルを各種記録媒体や通信を介して様々な電子機器に配信することで、それらの電子機器において、高精度な音声認識、画像認識、意図理解等を行うことが可能となる。さらに、上記実施の形態における標準モデル作成装置を各種電子機器に内蔵させることで、音声認識、画像認識、意図理解等の認識・認証機能を備えるスタンドアローンの電子機器を実現することもできる。

#### 【0220】

#### 【発明の効果】

以上の説明から明らかなように、本発明に係る標準モデル作成装置によれば、

2種類以上の参照モデルに対する標準モデルの確率または尤度を最大化または極大化するように標準モデルの統計量が計算され、標準モデルが作成されるので、音声データ等の入力データや教師データを必要とすることなく簡易に標準モデルが作成されるとともに、既に作成された複数の参照モデルを総合的に勘案した高精度な標準モデルが作成される。

#### 【0 2 2 1】

また、標準モデル作成装置の外部から新たな参照モデルを取り込み、取り込んだ参照モデルに基づいた標準モデルが作成されるので、様々な認識・認証対象に対応した汎用性の高い標準モデル作成装置が実現される。

#### 【0 2 2 2】

また、利用者の特徴、利用者の年齢、性別、利用環境などの利用情報に基づいて選択された参照モデルを基準として、標準モデルが作成されるので、認識・認証対象により特化した精度の高い標準モデルが作成される。

#### 【0 2 2 3】

また、通信路を介して送信されてきた参照モデルに基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識・認証システムの構築が実現される。

#### 【0 2 2 4】

また、標準モデルを使用する装置のCPUパワー、記憶容量、要求される認識精度、要求される認識処理時間などの仕様情報に基づいて標準モデルが作成されるので、特定の仕様条件を満たす標準モデルの生成が可能となり、計算エンジン等の認識・認証処理に必要なリソース環境に適した標準モデルの生成が実現される。

#### 【0 2 2 5】

また、通信路を介して送信されてきた仕様情報に基づいて標準モデルが作成されるので、遠隔制御による標準モデルの生成が可能になるとともに、通信システムを基盤とする認識・認証システムの構築が実現される。

#### 【0 2 2 6】

また、仕様情報に基づいて、標準モデルに含まれるガウス分布の混合数が動的

に決定されるので、認識・認証処理が実行される環境や要求仕様等に応じて標準モデルの構造を制御することが可能となる。たとえば、標準モデルを使用する認識装置のCPUパワーが小さい場合、記憶容量が小さい場合、要求される認識処理時間が短い場合などは標準モデルの混合数を少なく設定して仕様に合わせることができ、一方、要求される認識精度が高い場合などは混合数を多く設定して認識精度を高くすることができる。

#### 【0227】

また、混合数が異なる参照モデルに基づいて標準モデルが作成されるので、予め準備された多種多様な構造の参照モデルに基づく標準モデルの作成が可能となり、より認識・認証対象に適した精度の高い標準モデルの作成が実現される。

#### 【0228】

また、作成された標準モデルは内部の記憶手段に保持されるので、作成された標準モデルを一時的にバッファリングしておき、送信要求に対してすぐに出したり、他の装置に提供するデータサーバとしての役割を果たしたりすることが可能となる。

#### 【0229】

また、作成された標準モデルは空間的に離れた場所に置かれた外部装置に送信されるので、本標準モデル作成装置を標準モデル作成エンジンとして独立させたり、標準モデル作成装置を通信システムにおけるサーバとして機能させたりすることが可能になる。

#### 【0230】

また、準備される参照モデルの追加、更新等が行われるので、様々な認識・認証対象用のモデルを参照モデルとして追加したり、より精度の高い参照モデルに置き換えたりすることが可能となり、更新した参照モデルによる標準モデルの再生成や、生成された標準モデルを参照モデルとして再び標準モデルを作成するというフィードバックによる学習等が可能となる。

#### 【0231】

以上のように、本発明により、隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、



ベイジアンネットなどの確率モデルによる意図理解、確率モデルによるデータマイニングなどに用いる高精度な標準モデルが提供され、その実用的価値は極めて高い。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 2】

同サーバの動作手順を示すフローチャートである。

【図 3】

図 1 における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図 4】

図 2 におけるステップ S 1 0 1（標準モデルの作成）の詳細な手順を示すフローチャートである。

【図 5】

図 1 における第 1 近似部 1 0 4 による近似計算を説明する図である。

【図 6】

本発明の第 2 の実施の形態における標準モデル作成装置に係る S T B の全体構成を示すブロック図である。

【図 7】

同 S T B の動作手順を示すフローチャートである。

【図 8】

図 7 における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図 9】

図 7 における第 2 近似部による近似計算を説明する図である。

【図 1 0】

本発明の第 3 の実施の形態における標準モデル作成装置に係る P D A の全体構

成を示すブロック図である。

【図 1 1】

同 P D A の動作手順を示すフローチャートである。

【図 1 2】

図 1 0 における参照モデル記憶部に記憶されている参照モデルの例を示す図である。

【図 1 3】

同 P D A の選択画面の一例を示す。

【図 1 4】

図 1 0 における統計量推定部による統計量の推定手順を示す概念図である。

【図 1 5】

図 1 0 における第 3 近似部による近似計算を説明する図である。

【図 1 6】

本発明の第 4 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 1 7】

同サーバの動作手順を示すフローチャートである。

【図 1 8】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 1 9】

本発明の第 5 の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図 2 0】

同サーバの動作手順を示すフローチャートである。

【図 2 1】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図 2 2】

本発明の第6の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図23】

同サーバの動作手順を示すフローチャートである。

【図24】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図25】

本発明の第7の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図26】

同サーバの動作手順を示すフローチャートである。

【図27】

同サーバの動作手順を説明するための参照モデルおよび標準モデルの一例を示す図である。

【図28】

本発明の第8の実施の形態における標準モデル作成装置に係るサーバの全体構成を示すブロック図である。

【図29】

同サーバの動作手順を示すフローチャートである。

【図30】

本発明に係る標準モデル作成装置を具体的に適用したシステム例を示す図である。

【符号の説明】

101、401、501、601、701、801	サーバ
102、202、302、402、502、602、702	参照モデル
準備部	
103、203、303、403、503、603、703	参照モデル
記憶部	

104、206、306、406、506、606、706 標準モデル  
作成部

104a、206a、306a 標準モデル構造決定部

104b、206b、306b 初期標準モデル作成部

104c、206c、306c 統計量記憶部

104d、206d、306d 統計量推定部

104e 第1近似部

111、311、511、611、711 読み込み部

112、413、513 書き込み部

201 STB

204、304 利用情報作成部

205、305、405、505、605、705 参照モデル選択部

206e 第2近似部

211、312 マイク

212、512 音声データ蓄積部

213 音声認識部

301 PDA

306e 第3近似部

307、607 仕様情報作成部

313 雑音識別部

404、504、604、704 利用情報受信部

411 カメラ

412 画像データ蓄積部

414 電話

507、707 仕様情報受信部

512、614、712 端末装置

708 標準モデル記憶部

709 標準モデル送信部

712a 携帯電話機

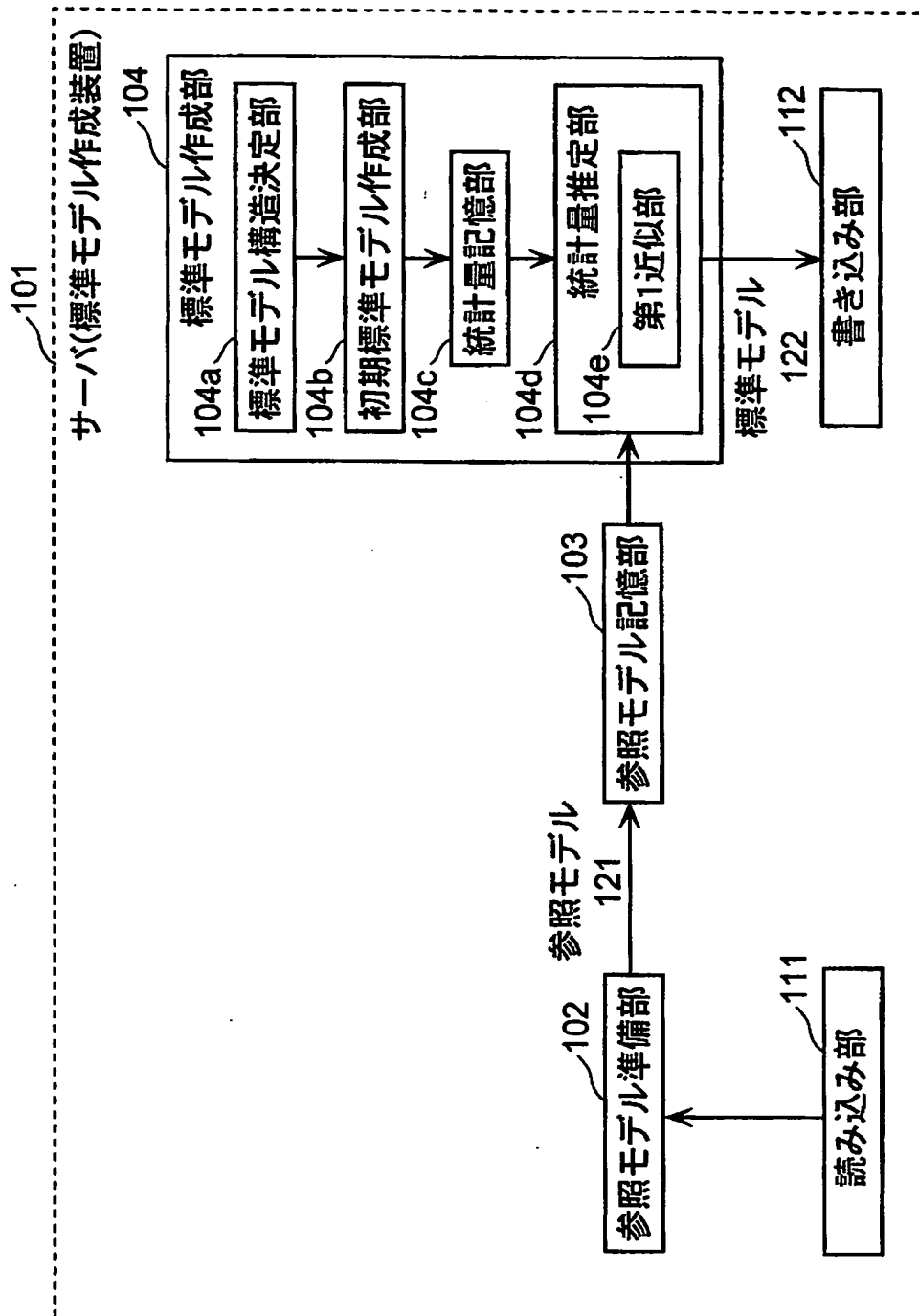
7 1 2 b カーナビゲーション装置

8 1 0 参照モデル受信部

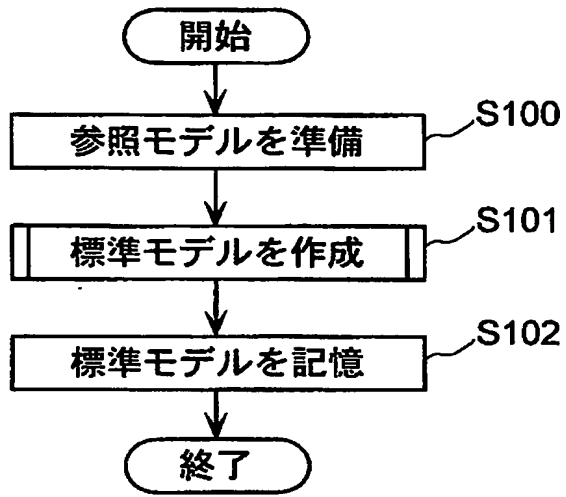
【書類名】

図面

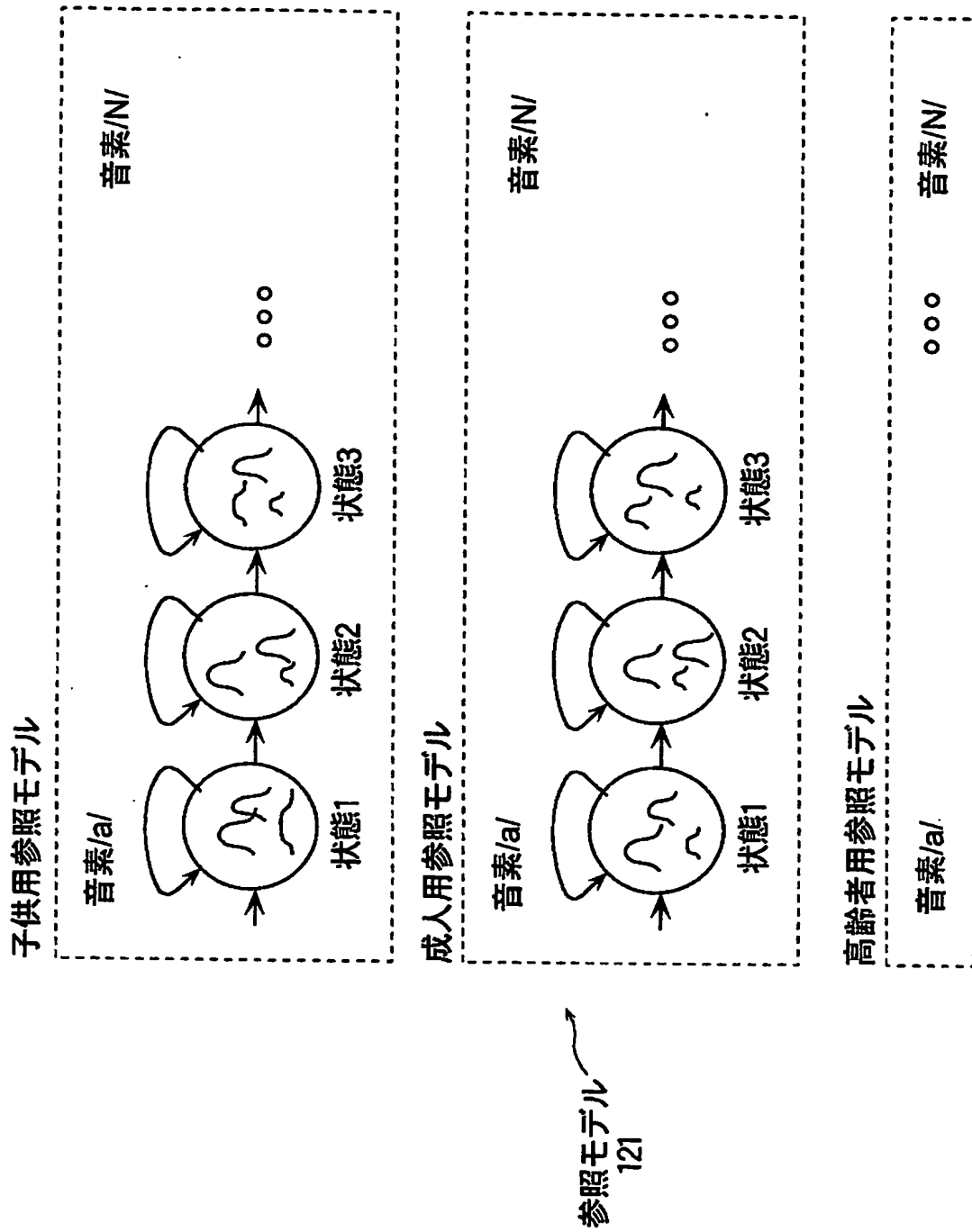
【図1】



【図 2】

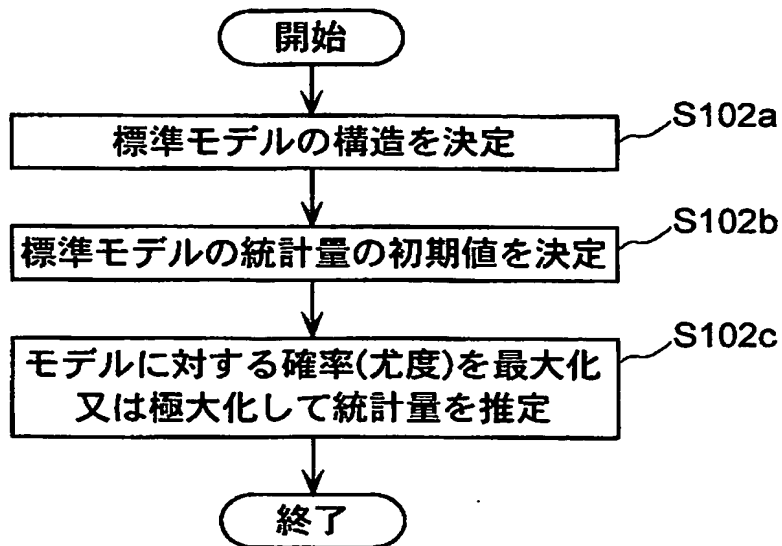


【図 3】

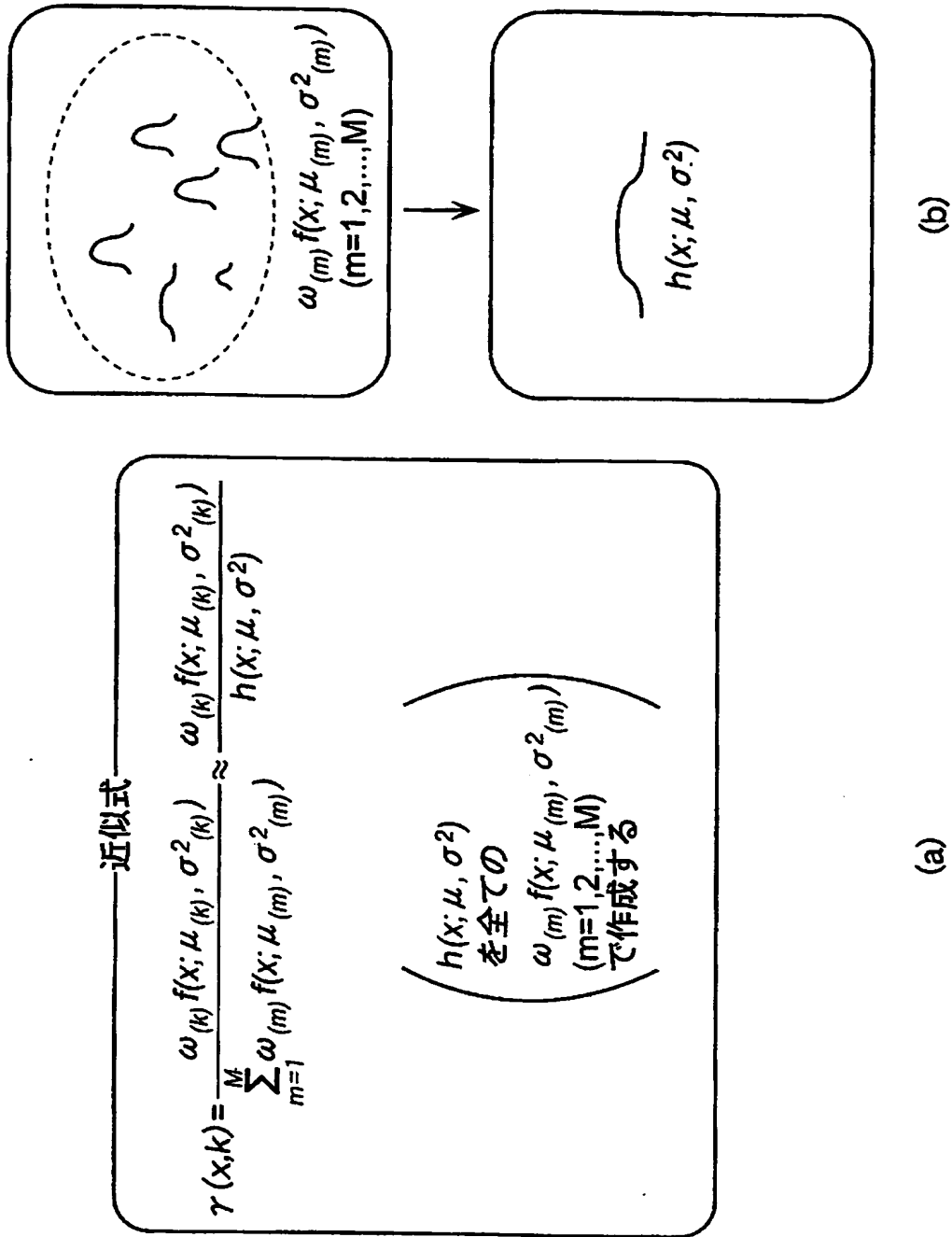




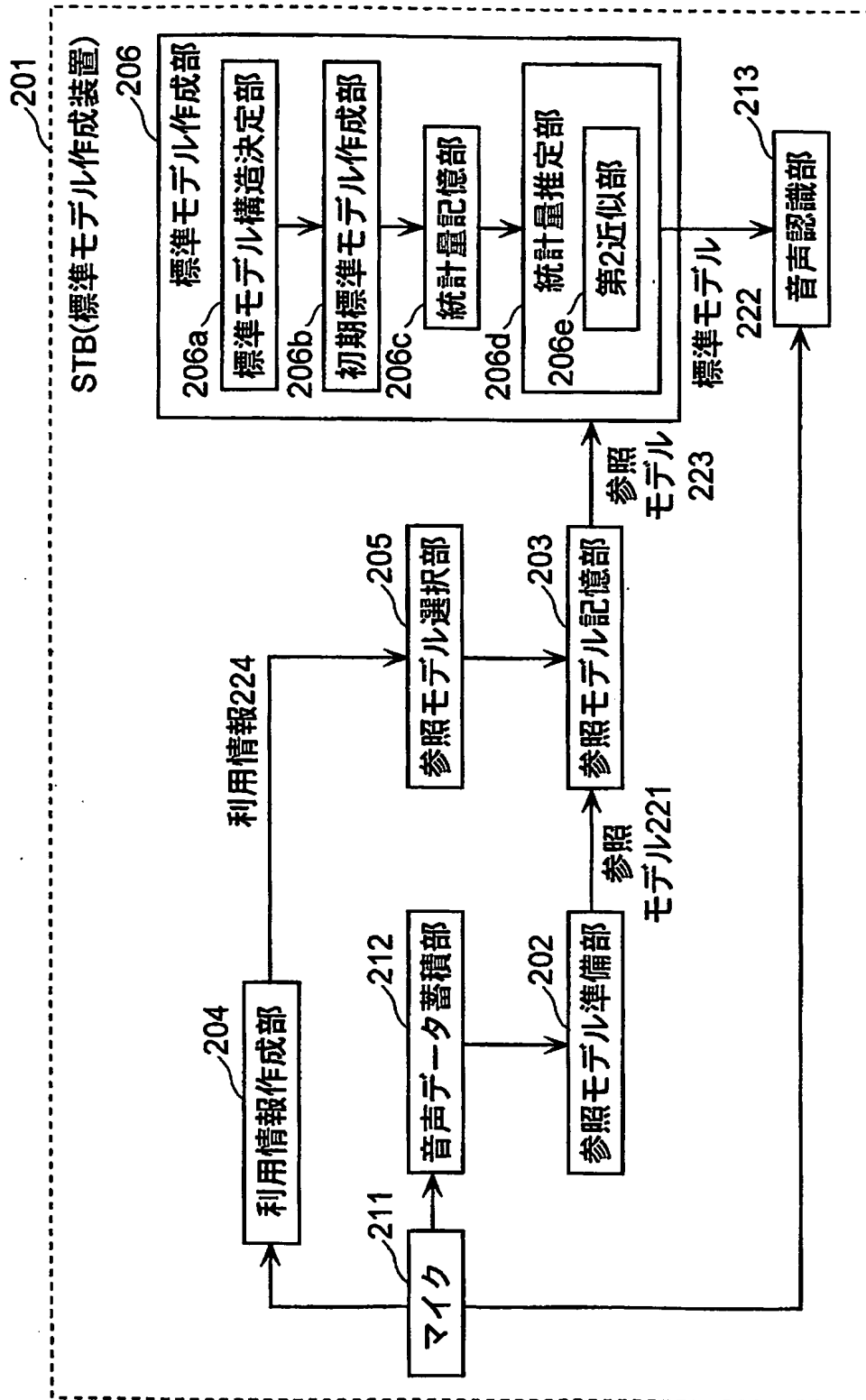
【図 4】



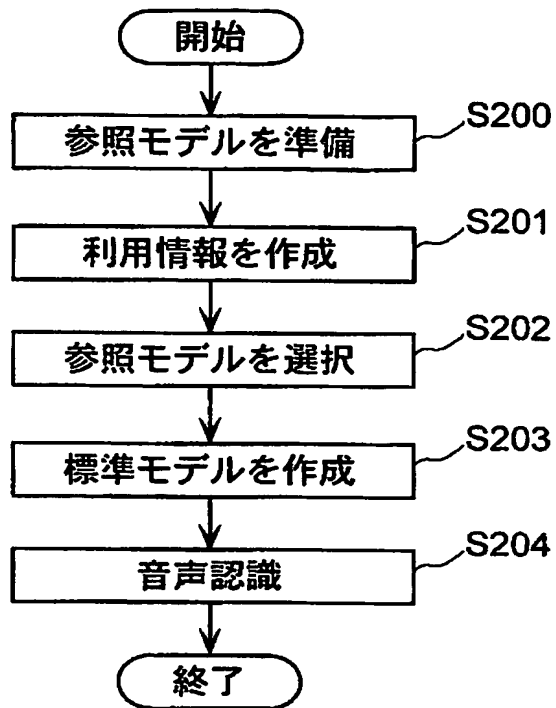
【図 5】



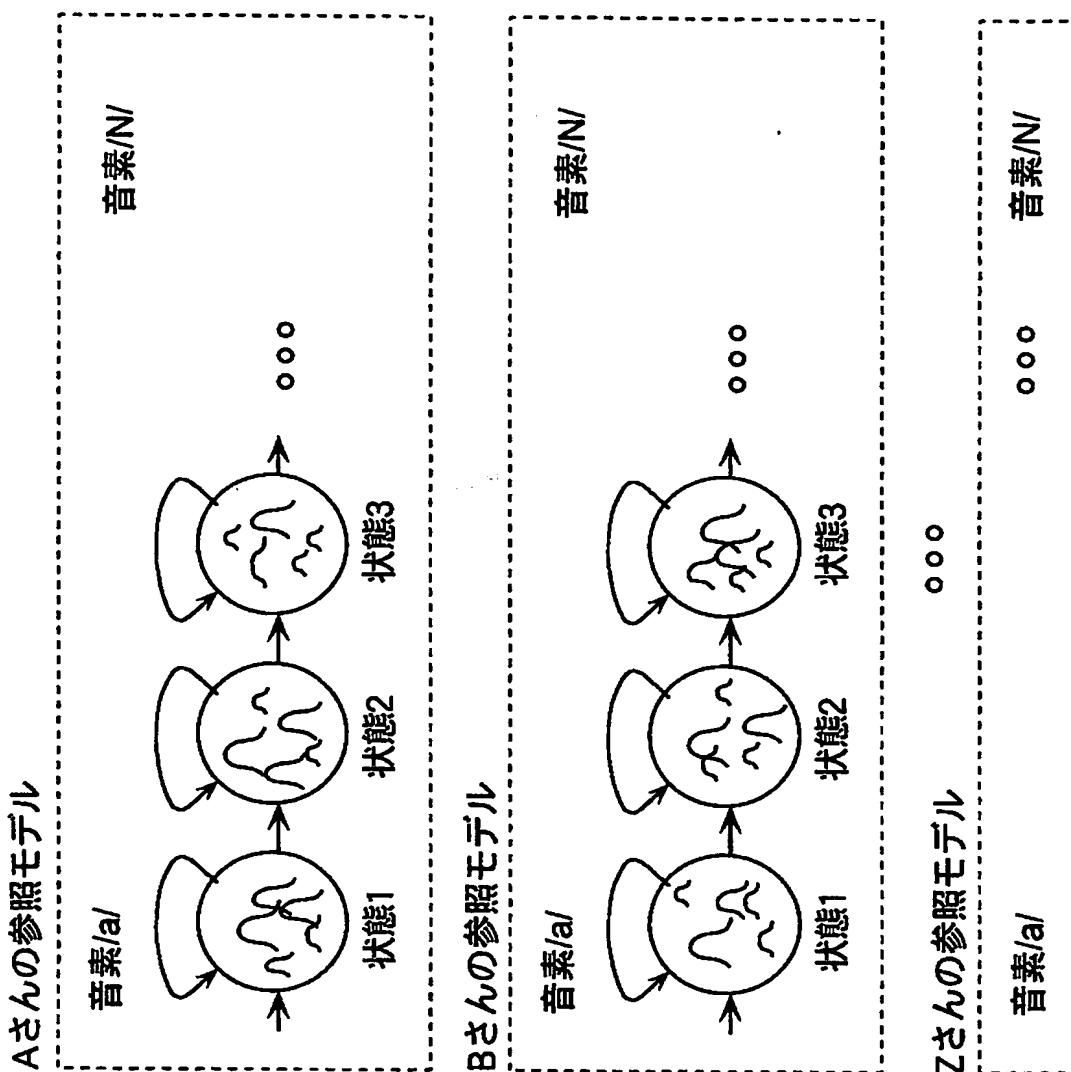
【図 6】



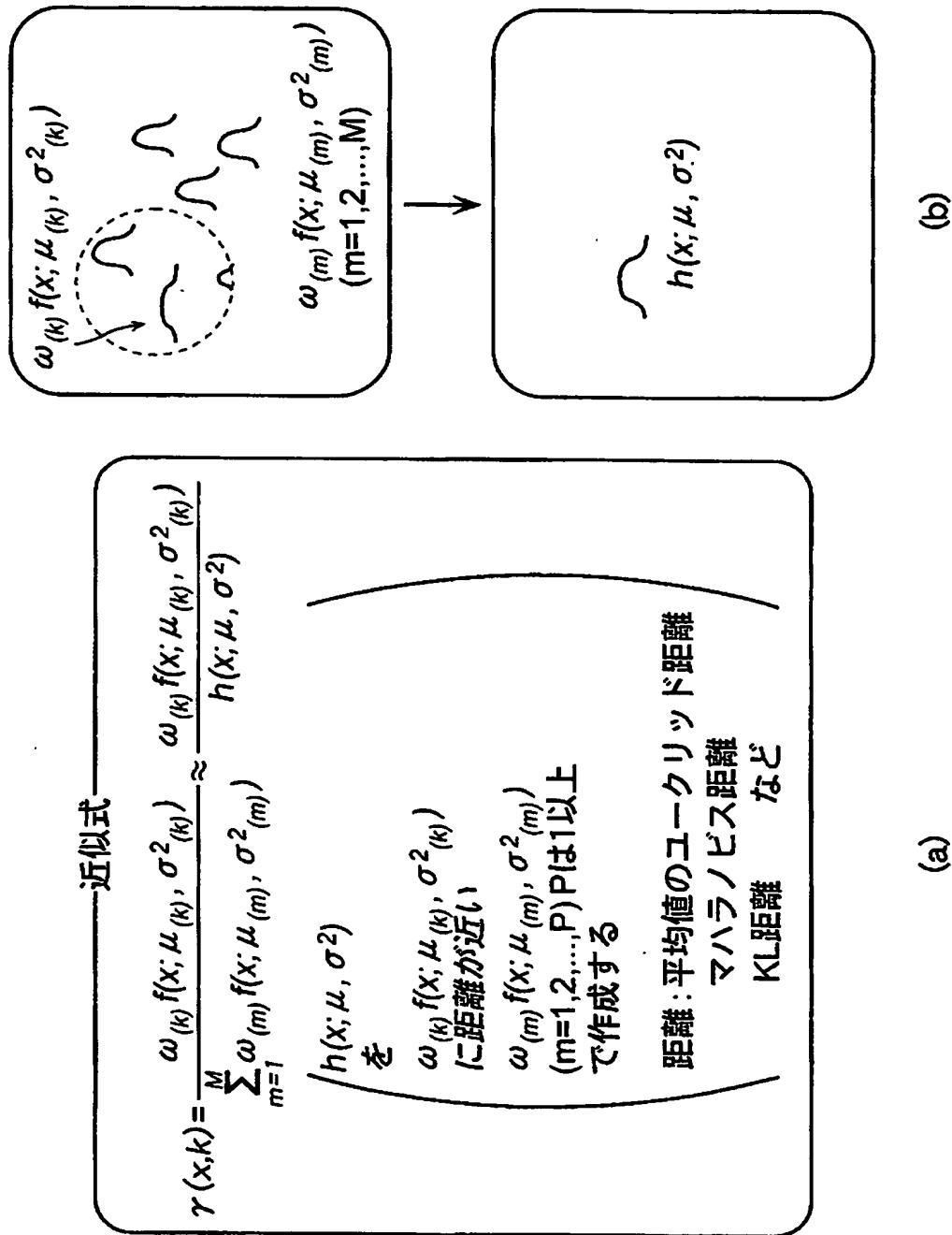
【図 7】



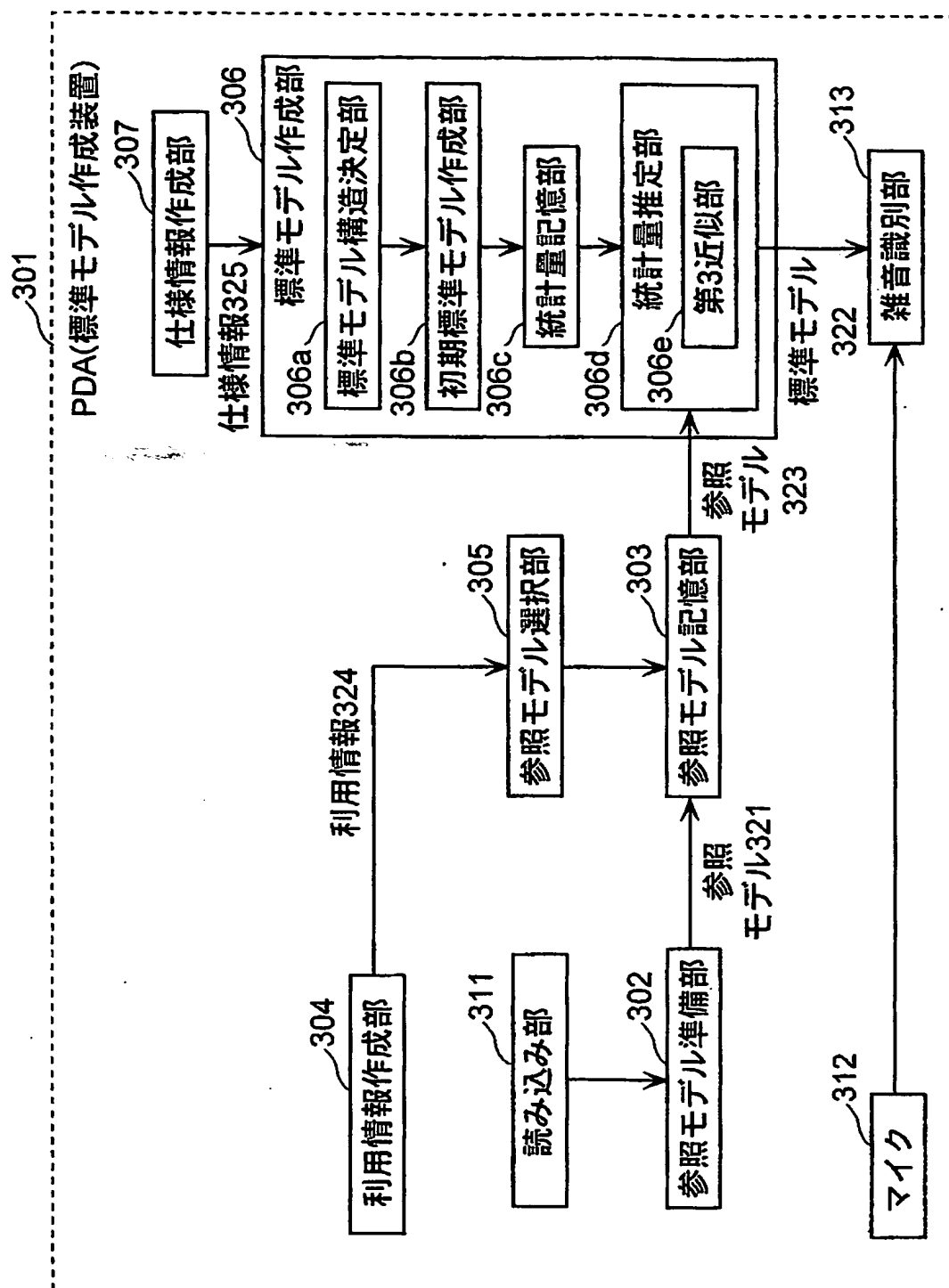
【図 8】



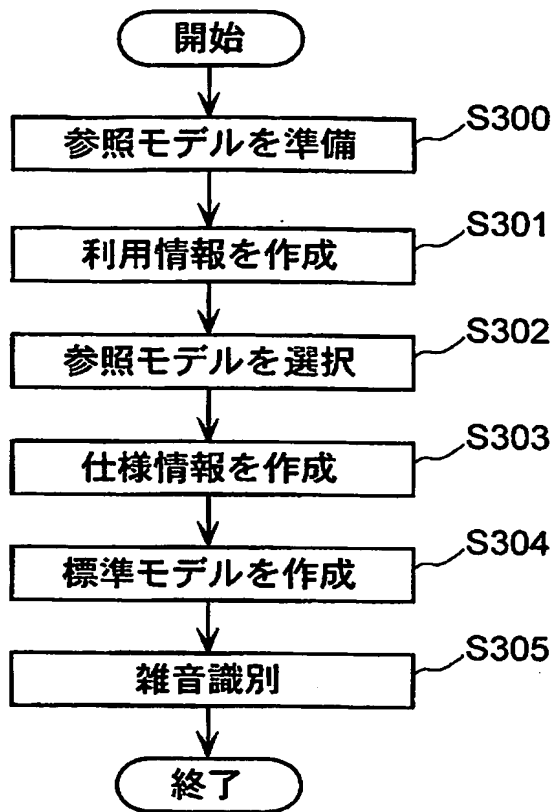
【図 9】



【図10】

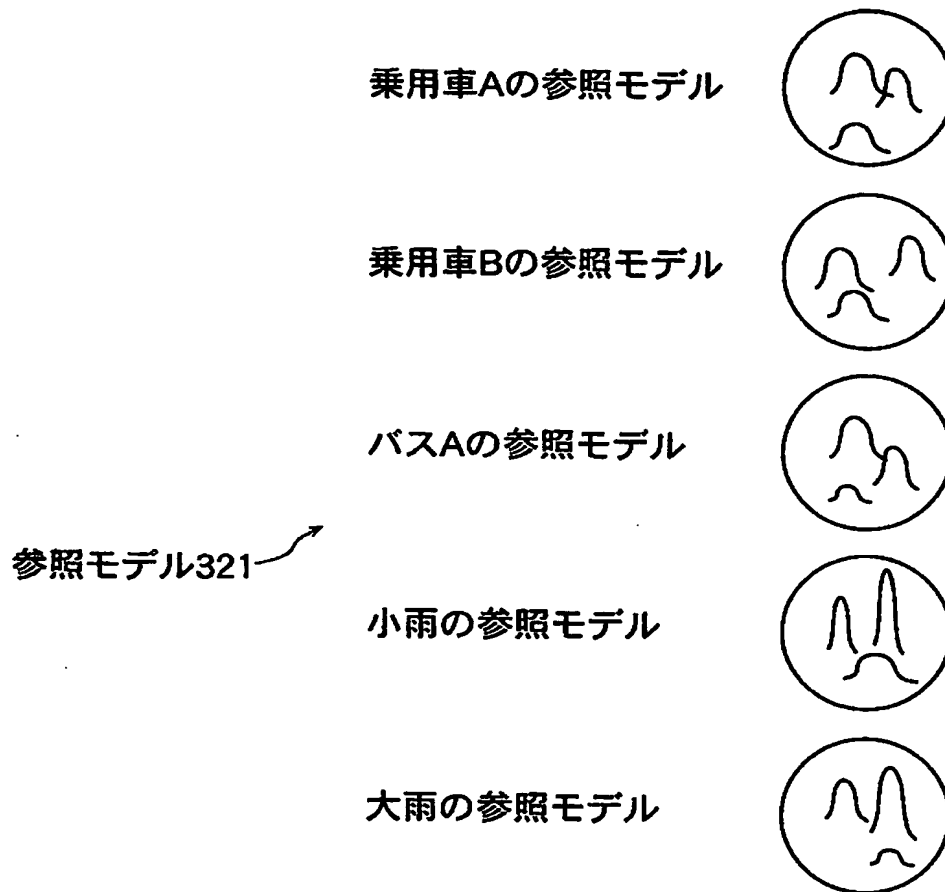


【図 11】





【図 12】



【図 13】

識別する雑音の種類

1. 自動車

1.1. 乗用車

1.2. バス

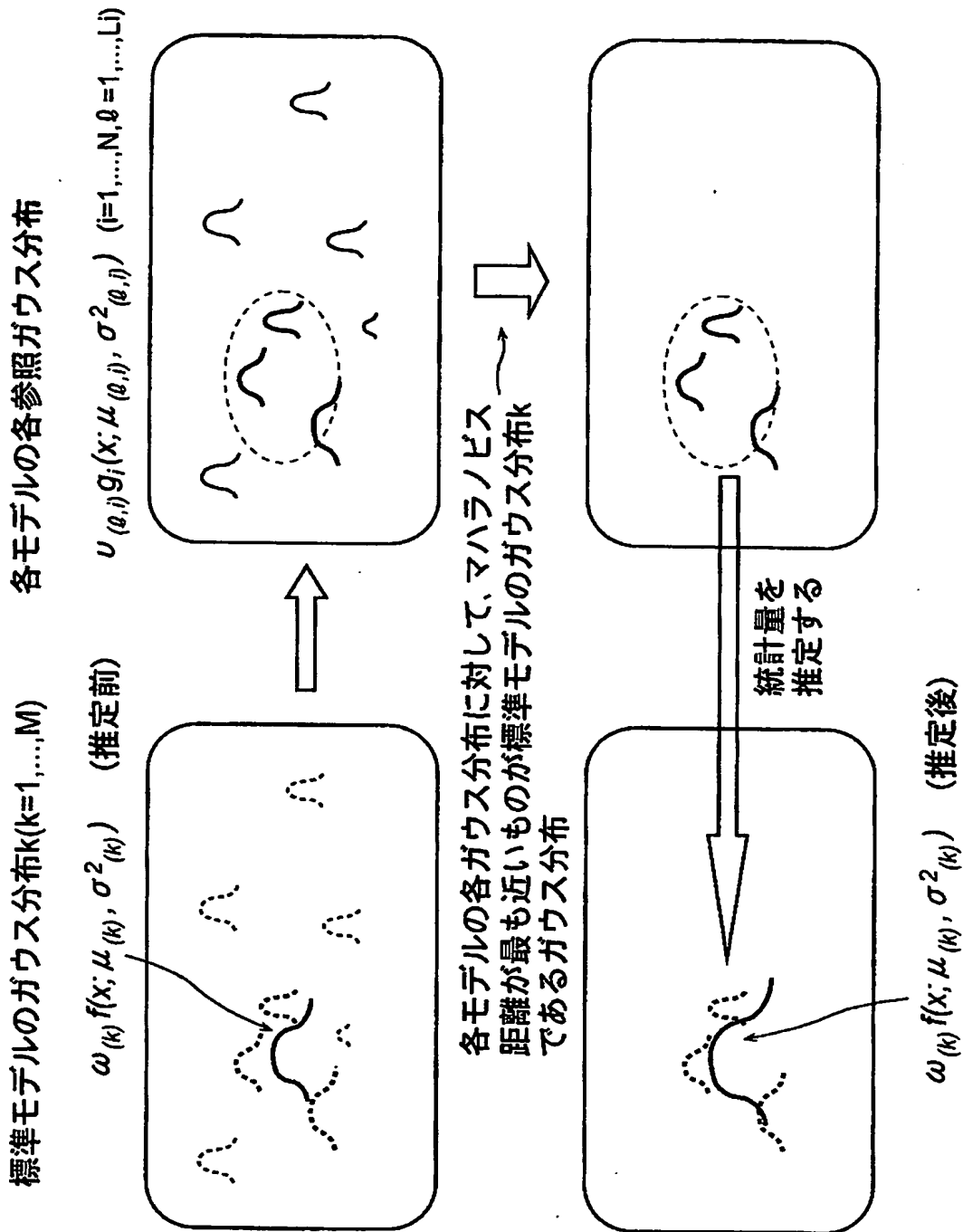
1.3. トラック

2. 雨

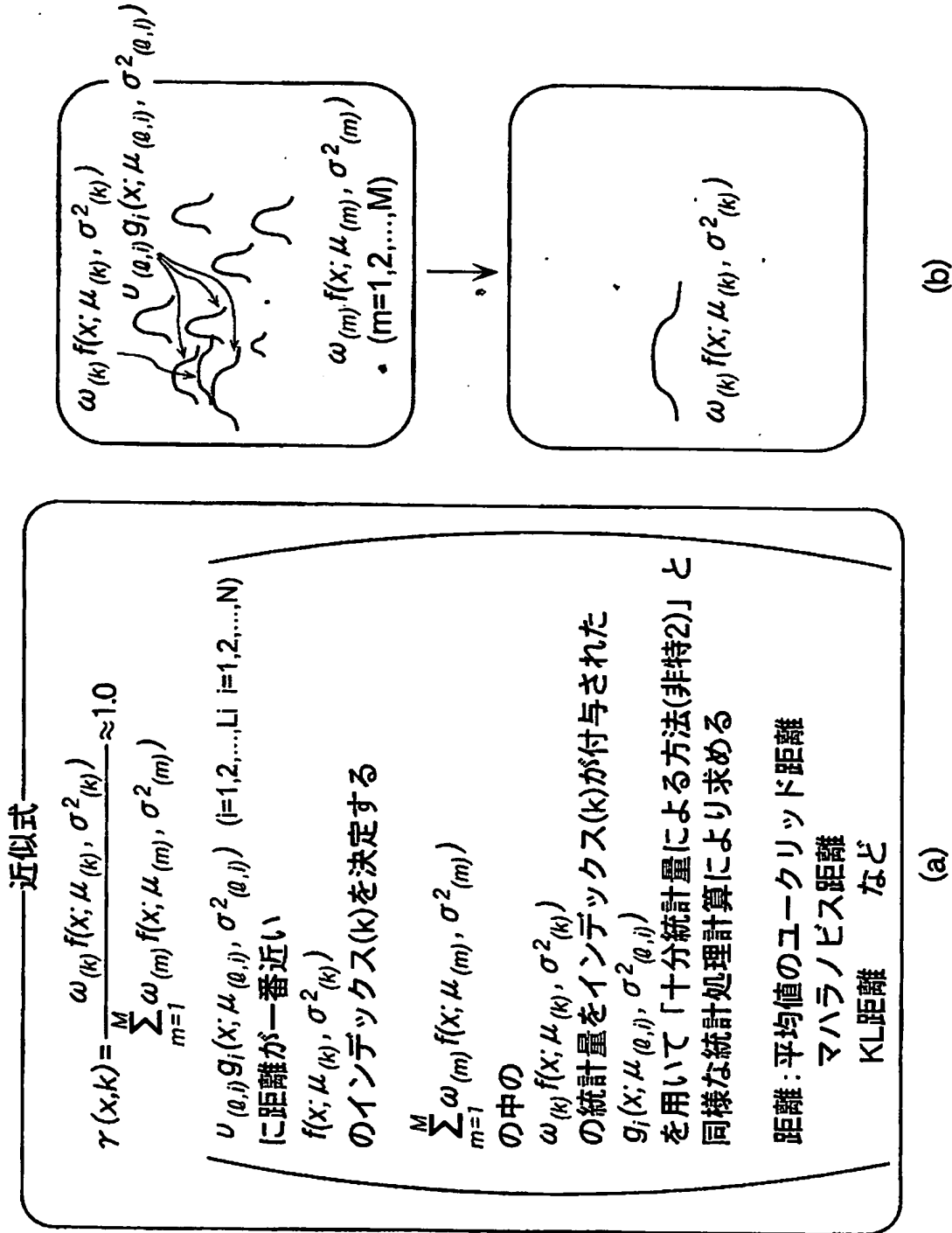
3. 飛行機

4. 警報音

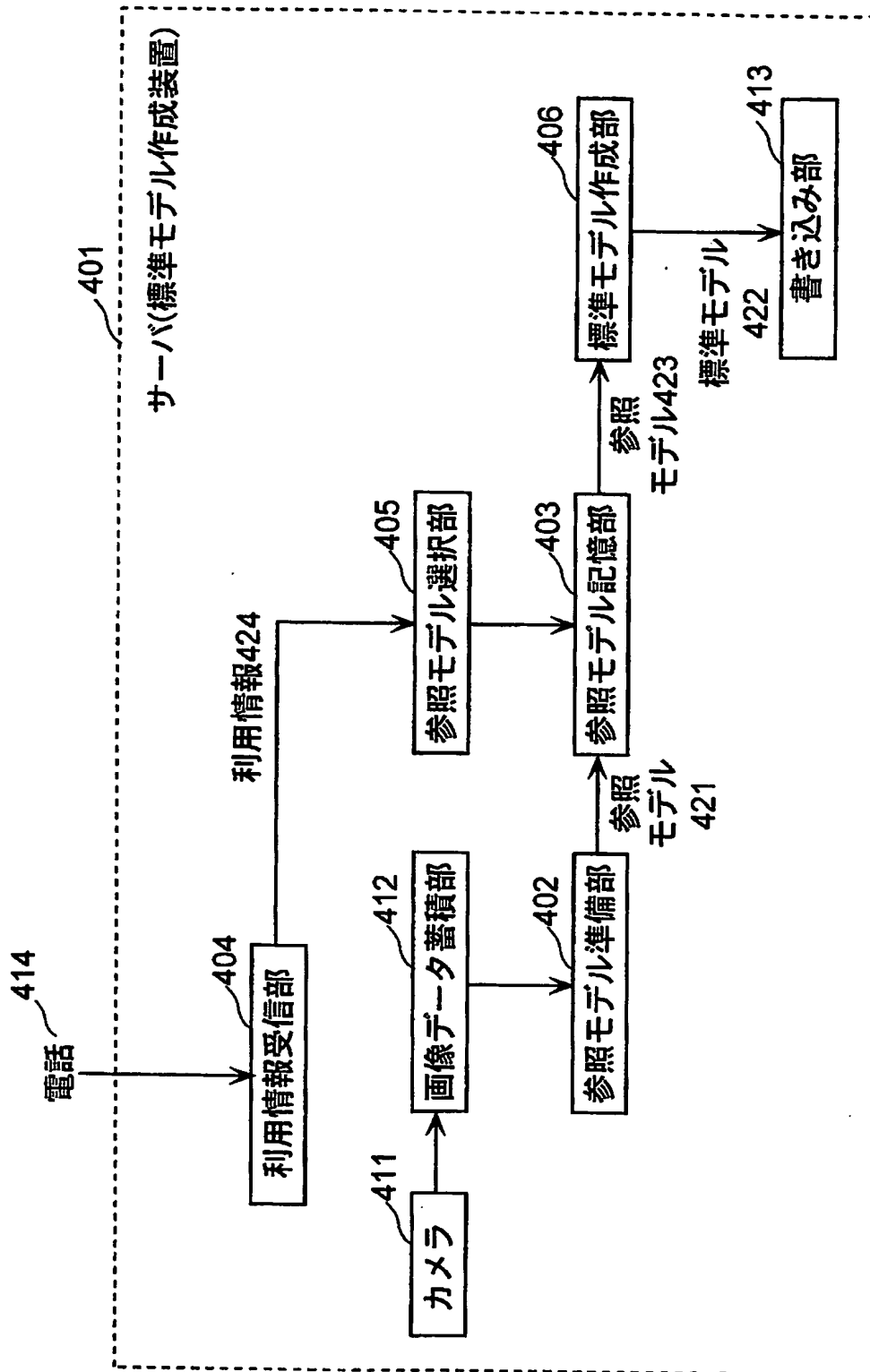
【図 14】



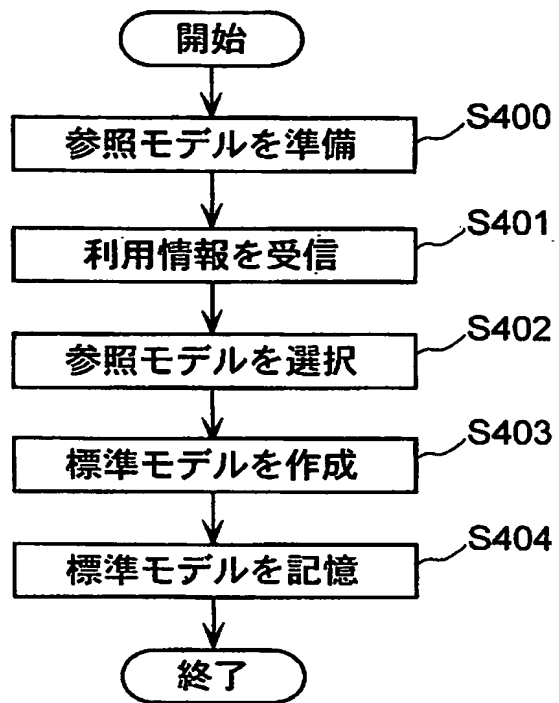
【図 15】



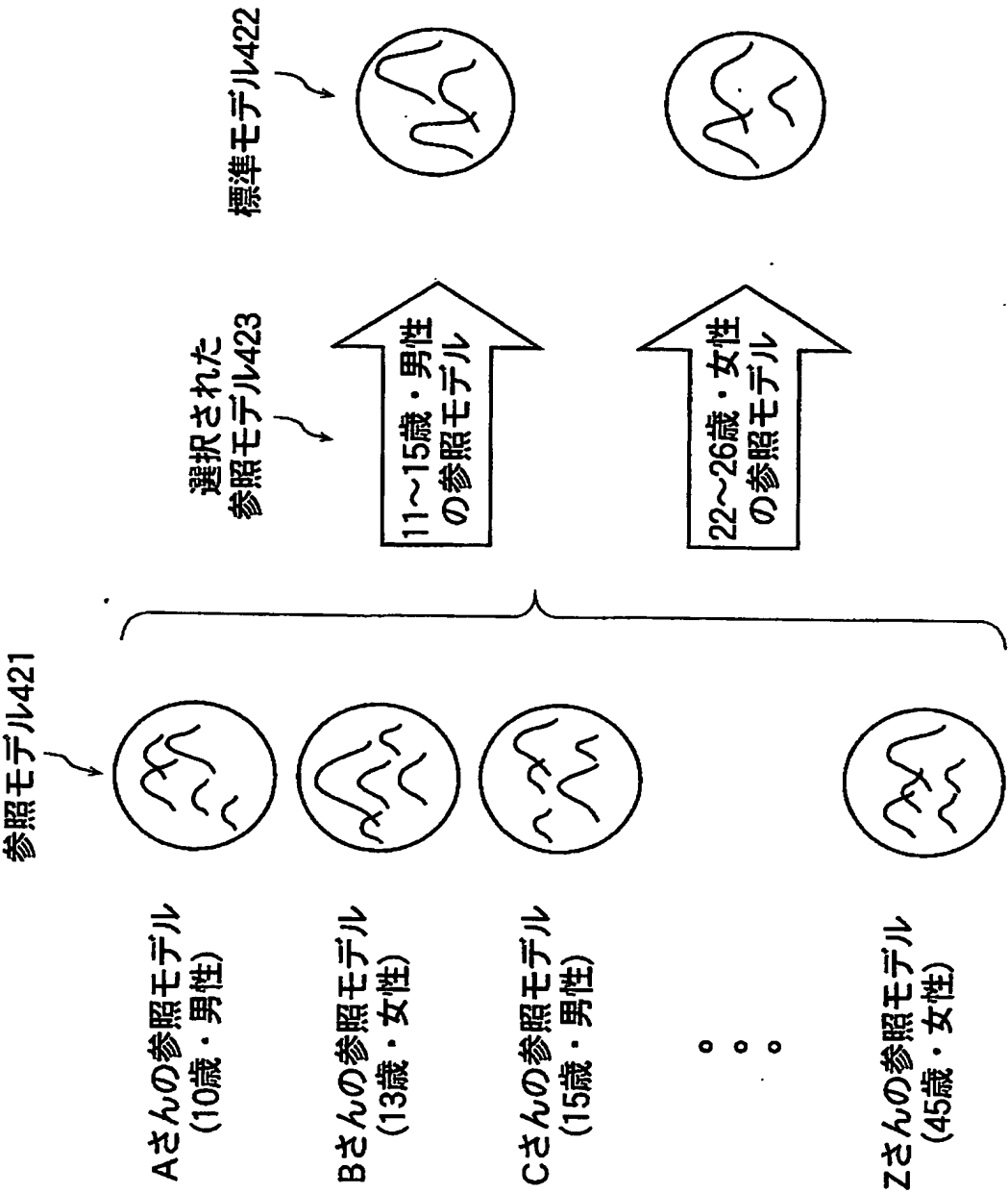
【図 16】



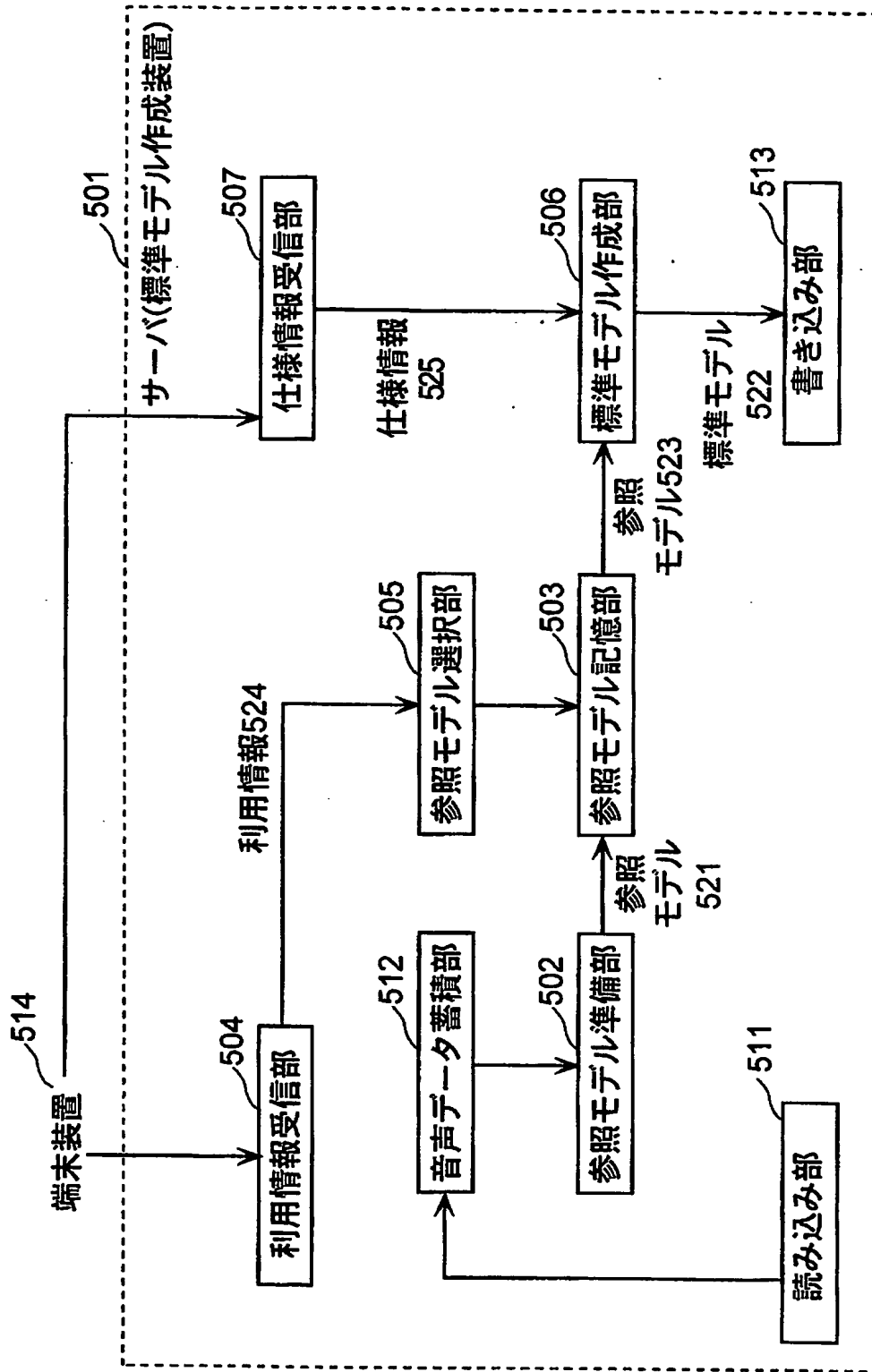
【図 17】



【図 18】

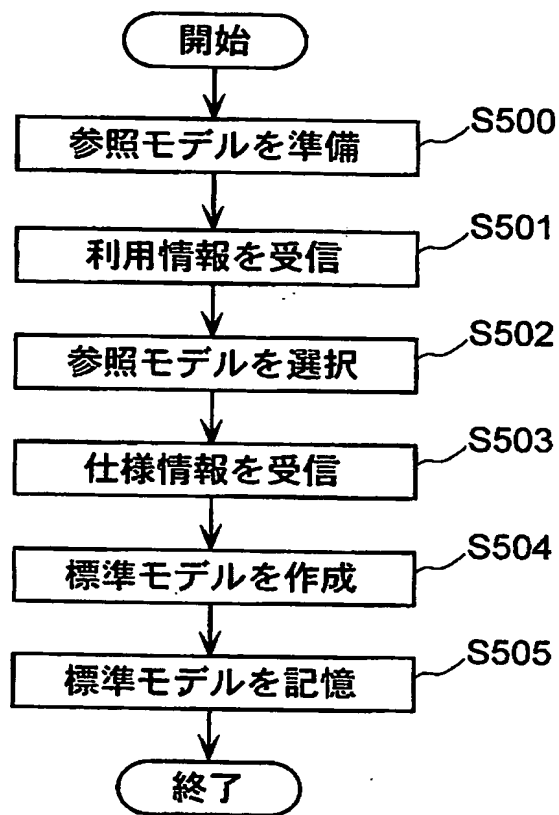


【図 19】

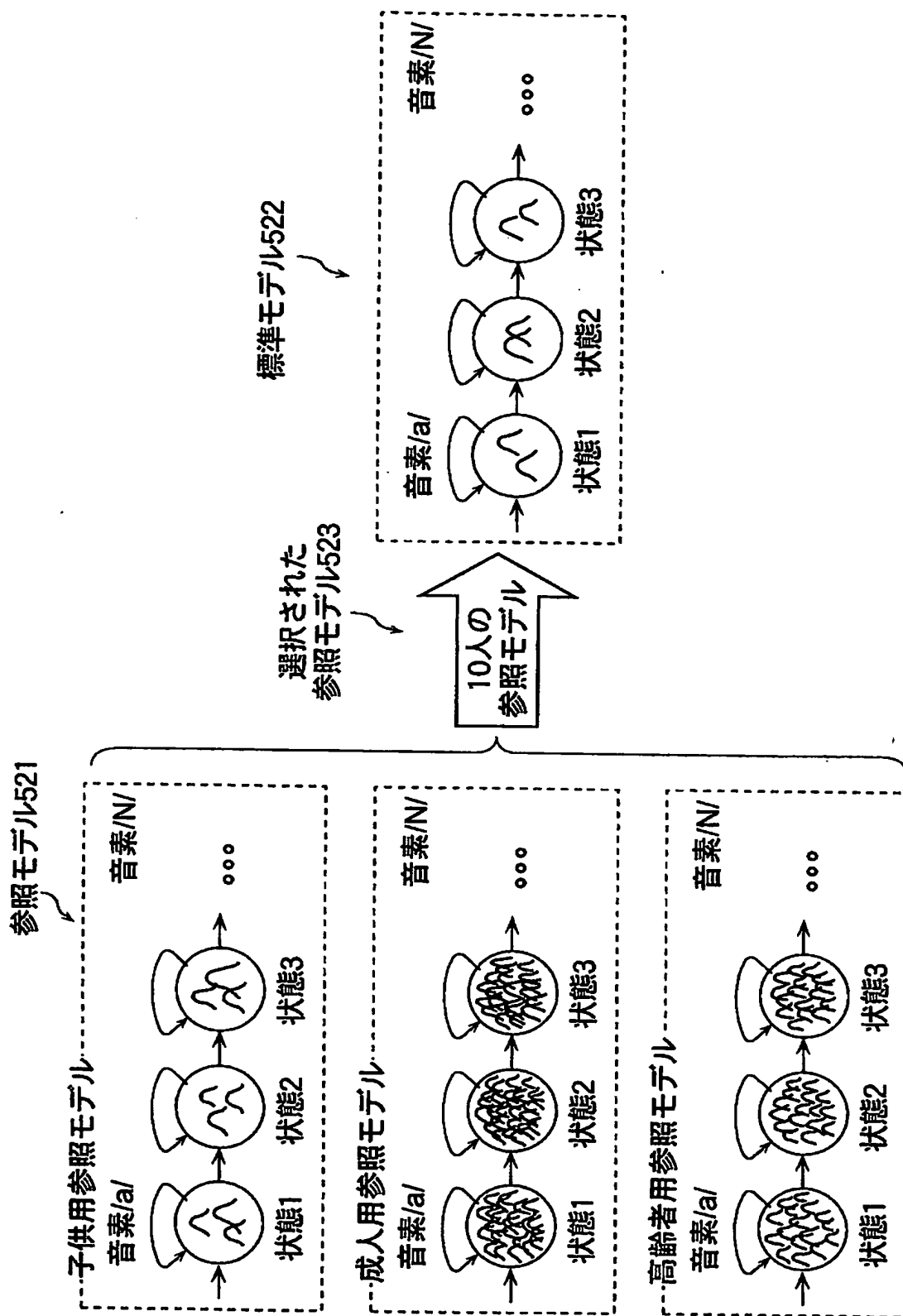




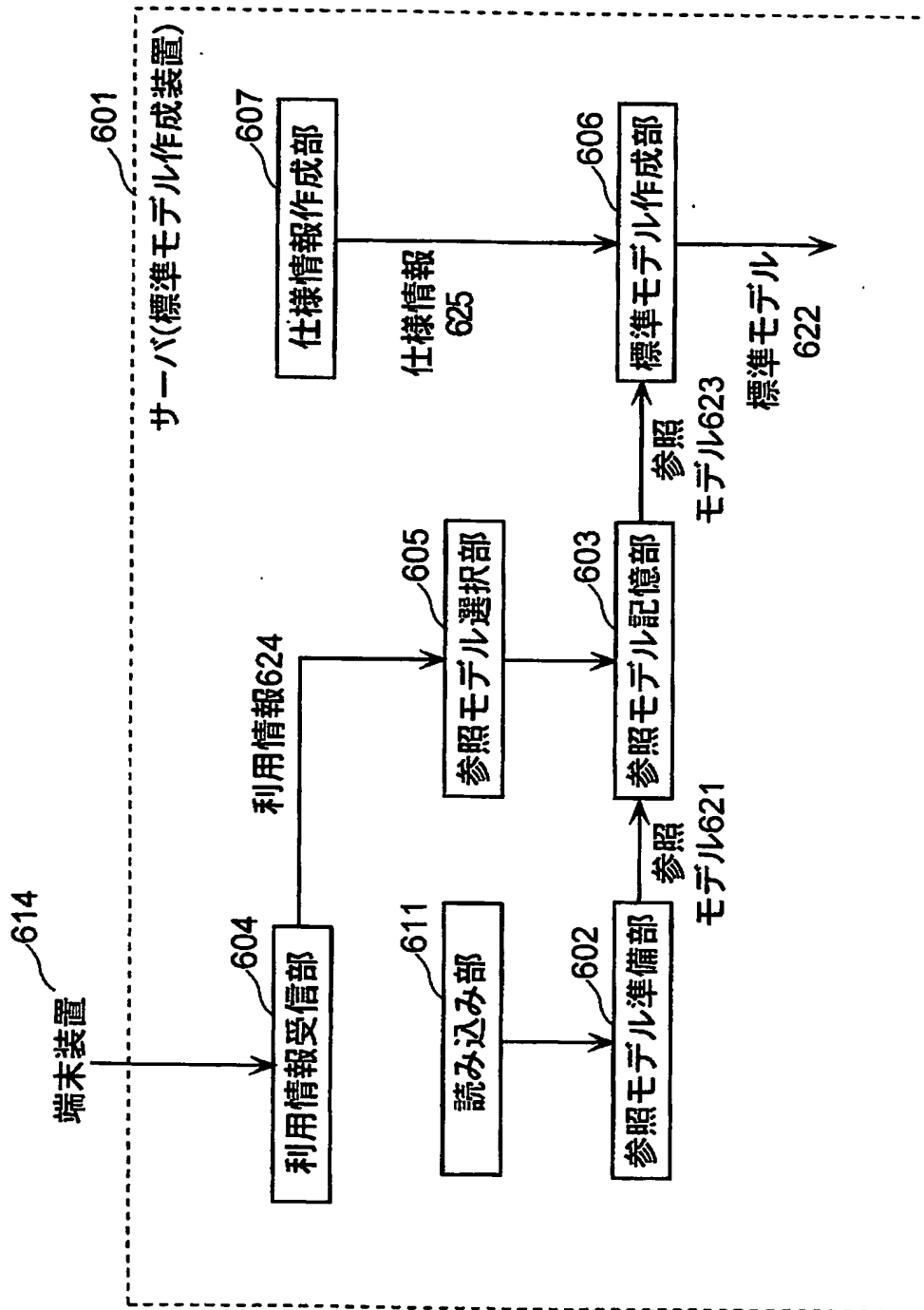
【図 20】



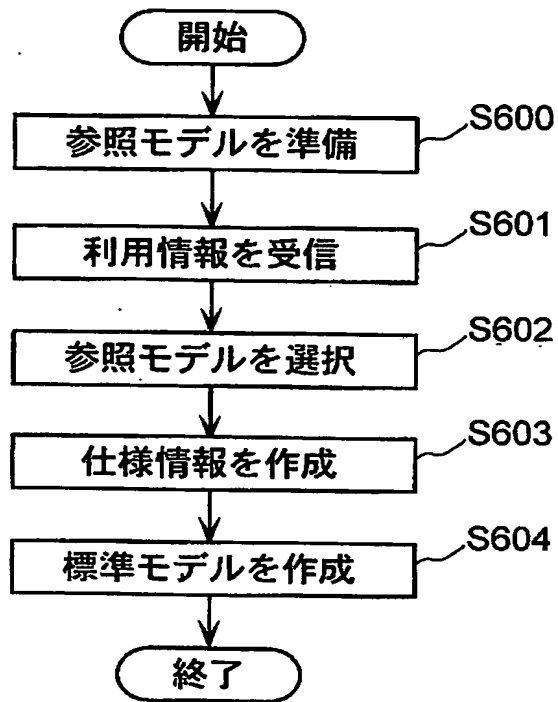
【図 21】



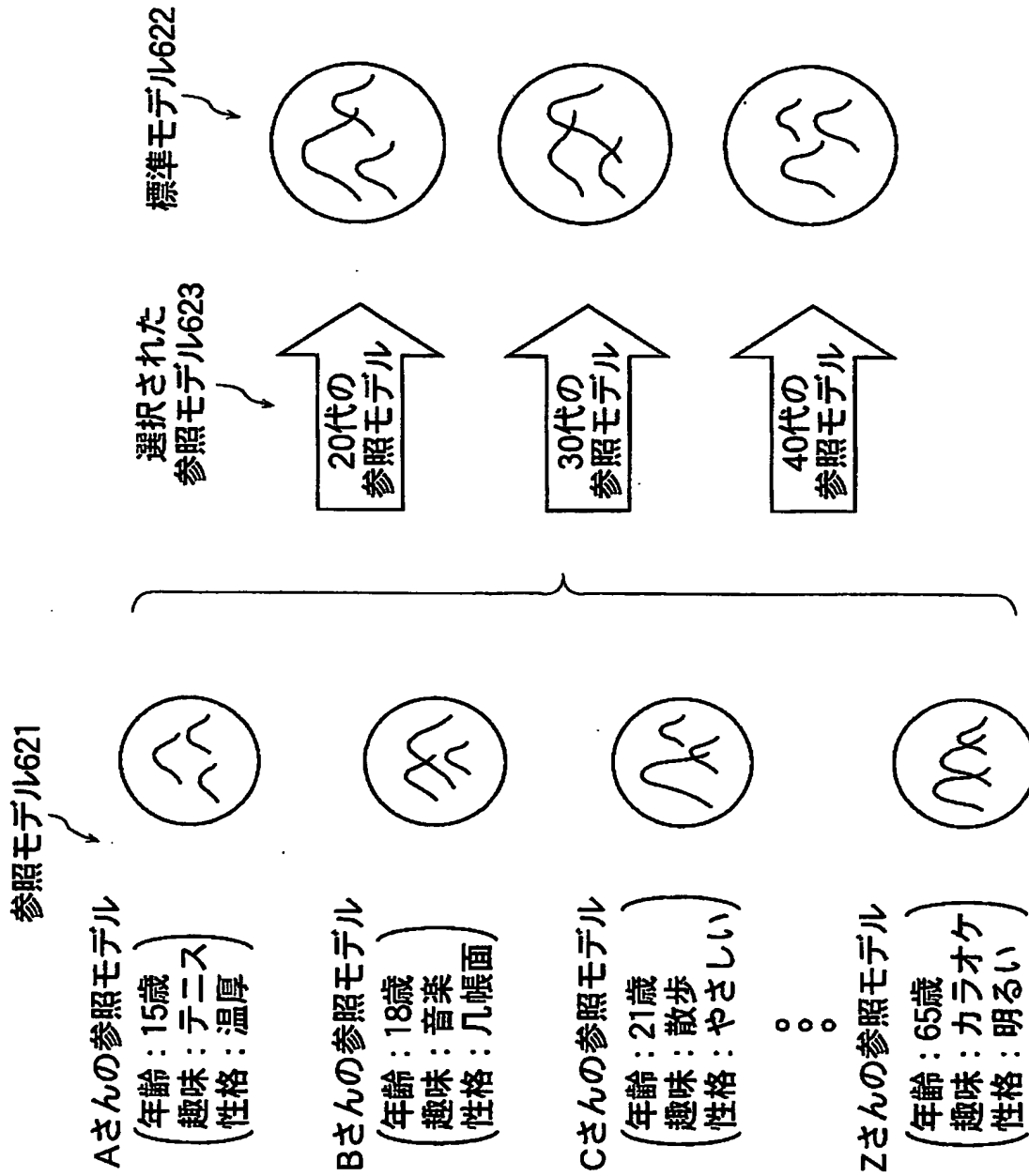
【図 22】



【図 23】

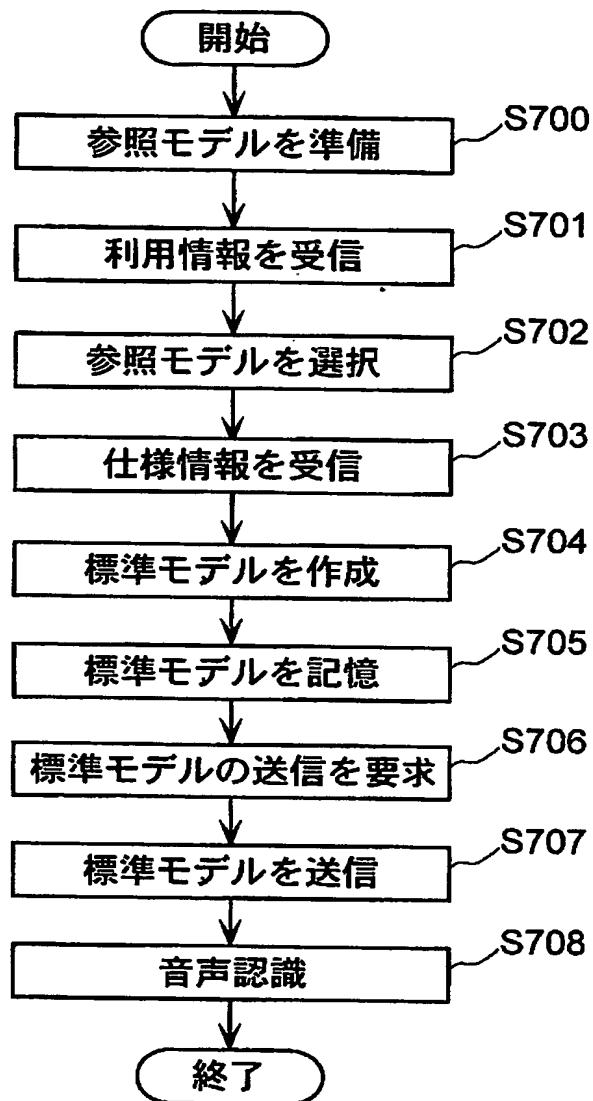


【図 24】

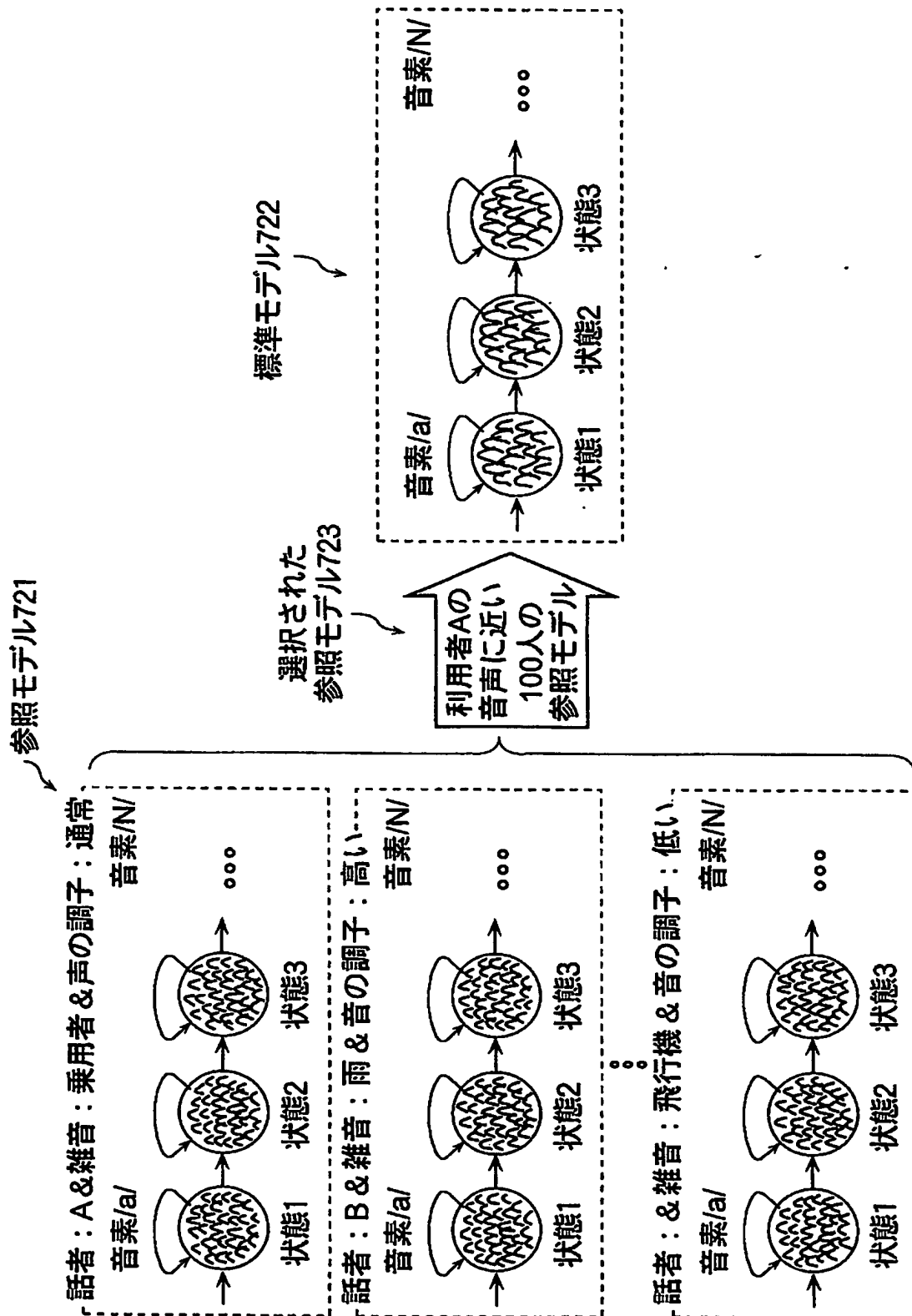




【図 26】

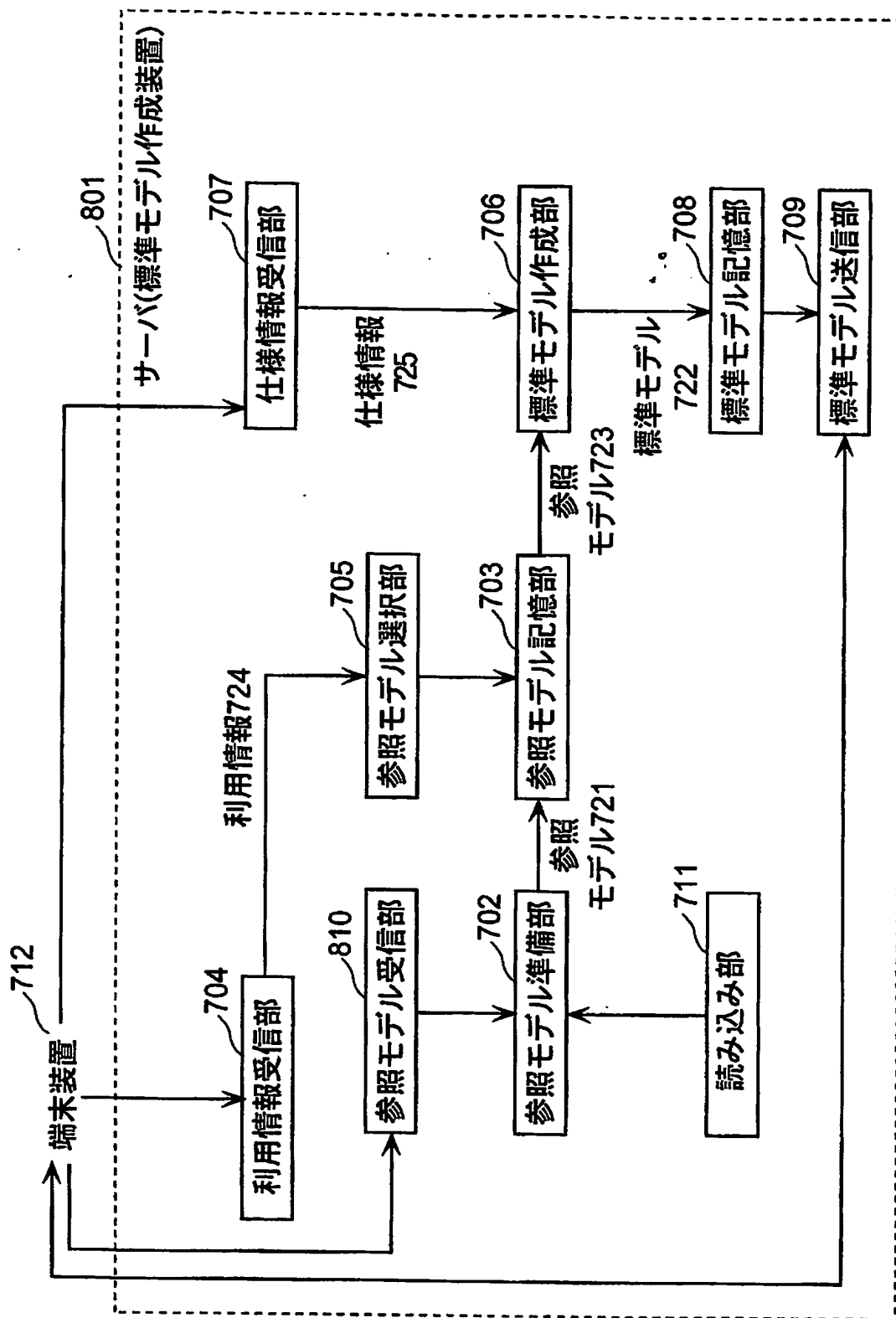


【図 27】

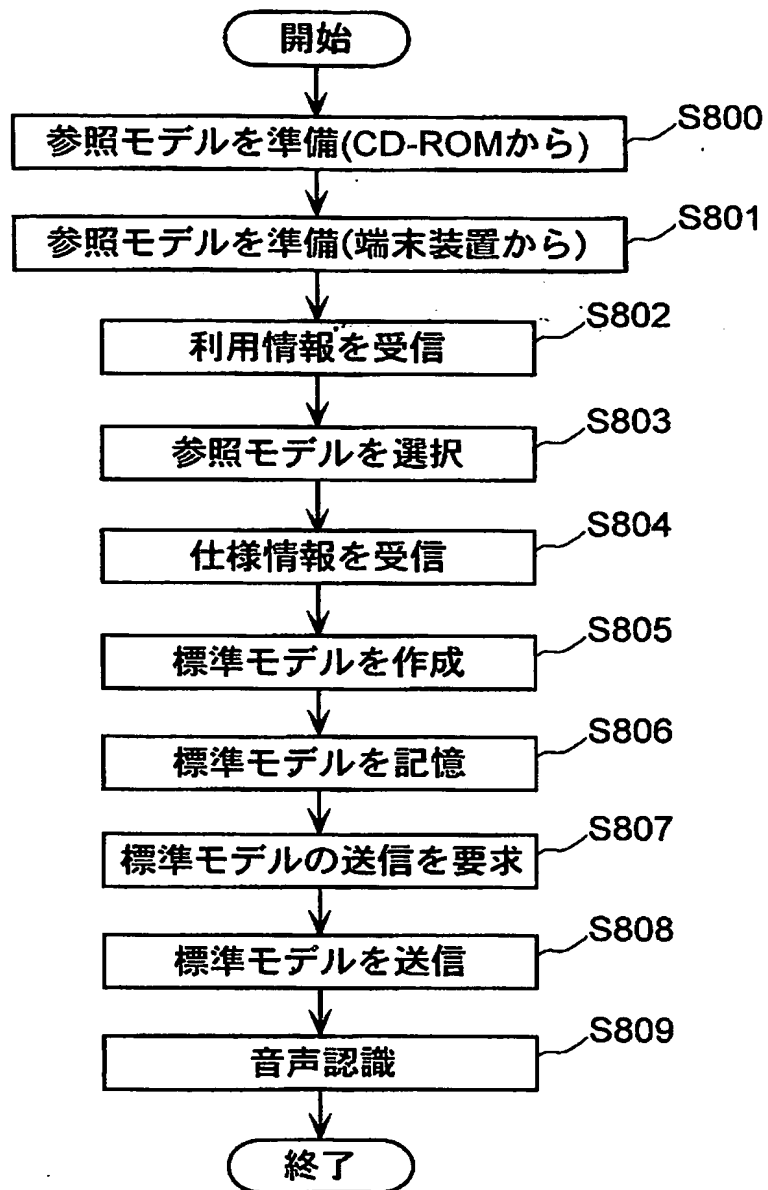




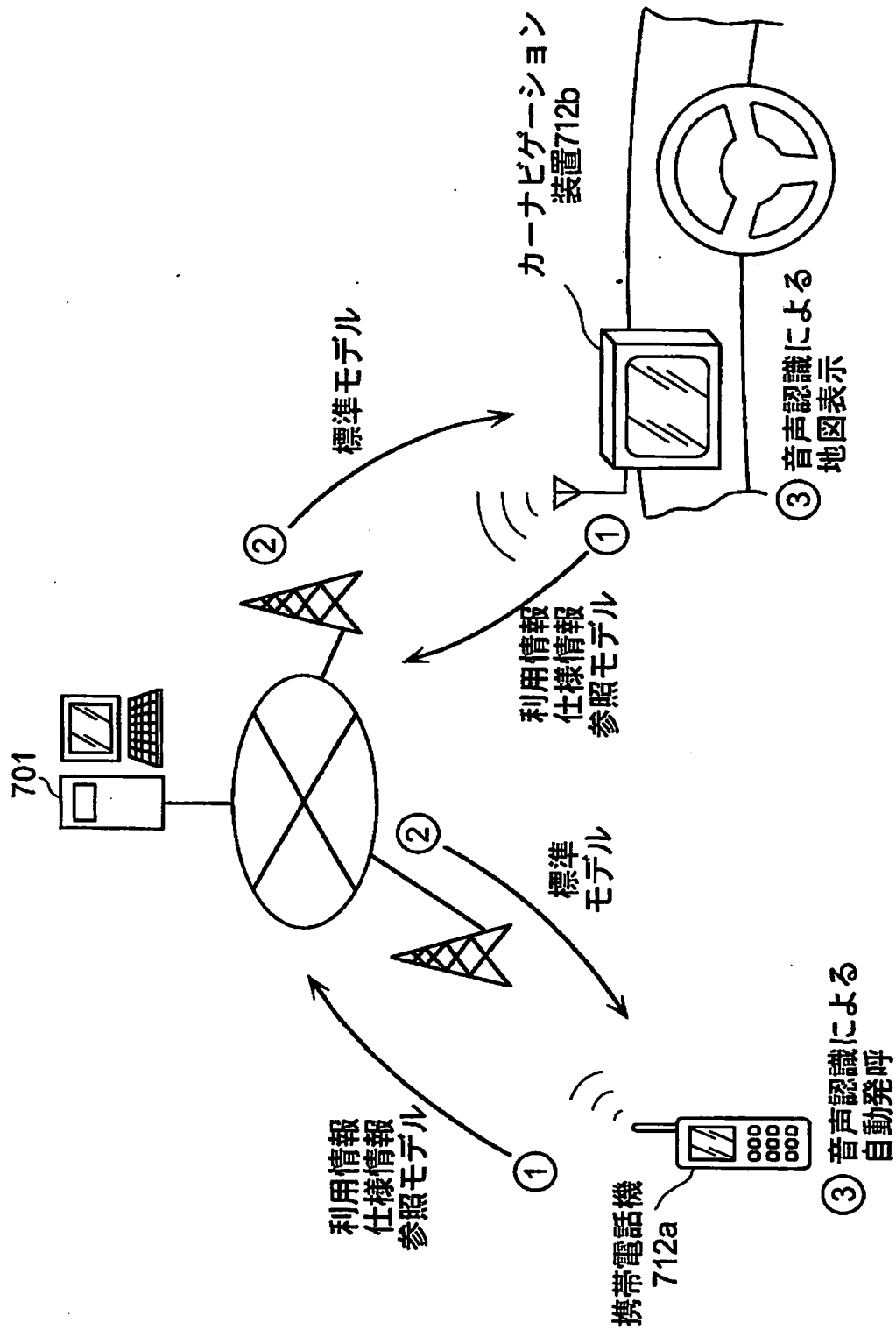
【図 28】



【図 29】



【図 30】



【書類名】 要約書

【要約】

【課題】 隠れマルコフモデル、ベイズ理論、線形判別分析などの確率モデルによる音声認識、文字認識、画像認識などのパターン認識、ベイジアンネットなどの確率モデルによる意図理解、確率モデルによるデータマイニングなどに用いる高精度な標準モデルを提供する。

【解決手段】 2種類以上の参照モデルを準備する参照モデル準備部102と、参照モデル準備部102が準備した参照モデル121を記憶する参照モデル記憶部103と、参照モデル記憶部103が記憶している2種類以上の参照モデルに対する確率または尤度を最大化または極大化するように標準モデルの統計量を計算して標準モデル122を作成する標準モデル作成部104とを備える。

【選択図】 図1

認定・付加情報

特許出願の番号	特願 2002-338652
受付番号	50201762756
書類名	特許願
担当官	第八担当上席 0097
作成日	平成14年11月22日

<認定情報・付加情報>

【提出日】	平成14年11月21日
-------	-------------

次頁無

特願 2 0 0 2 - 3 3 8 6 5 2

出 願 人 履 歴 情 報

識別番号

[ 0 0 0 0 0 5 8 2 1 ]

1. 変更年月日

1 9 9 0 年 8 月 2 8 日

[変更理由]

新規登録

住 所

大阪府門真市大字門真 1 0 0 6 番地

氏 名

松下電器産業株式会社